Biochemical and structural characterization of the ATP-dependent maturation factor of acetyl-CoA synthase

D i s s e r t a t i o n

zur Erlangung des akademischen Grades

d o c t o r r e r u m n a t u r a l i u m

(Dr. rer. nat.)

im Fach Biologie

eingereicht an der

Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von

Diplom-Chemikerin Christina Maria Gregg

geb. Wörmann

Präsidentin der Humboldt-Universität zu Berlin:

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät:

Prof. Dr. Bernhard Grimm

Gutachter:

1. Prof. Dr. Holger Dobbek 2. Prof. Dr. Thomas Eitinger 3. Prof. Dr. Peter Hildebrandt

Tag der mündlichen Prüfung: 05.04.2017

Diese Arbeit wurde von April 2012 bis Juli 2016 an der Humboldt Universität zu Berlin in der Arbeitsgruppe Biochemie/Strukturbiologie unter Leitung von Prof. Holger Dobbek angefertigt. Teile der im zeitlichen Rahmen dieser Dissertation erzielten Ergebnisse sind in folgenden Publikationen veröffentlicht:

This research was carried out between April 2012 and July 2016 at the Humboldt University of Berlin, in the biochemistry/structural biology research group under the supervision of Prof. Holger Dobbek. Some results of this research have been published in the following publications:

C. M. Gregg, S. Goetzl, J.-H. Jeoung, and H. Dobbek. AcsF catalyzes the ATP-dependent insertion of Nickel into the Ni,Ni-[4Fe4S] cluster of acetyl-CoA synthase. *J. Biol. Chem.* 291(35):18129-18138, 2016.

J.-H. Jeoung, S. Goetzl, S. E. Hennig, J. Fesseler, C. Wörmann, J. Dendra and H. Dobbek. The extended reductive acetyl-CoA pathway: ATPases in metal cluster maturation and reductive activation. *Biol. Chem.* 395:545-58, 2014.

Zusammenfassung

Acetyl-CoA Synthase (ACS) ist ein Schlüsselenzym des reduktiven Acetyl-CoA-Weges, das die Reaktion von CO, CoA und einem Methylkation zu Acetyl-CoA katalysiert. Die Reaktion läuft an einem Ni,Ni-$[Fe_4S_4]$-Cluster (A-Cluster) ab, in dem zwei Nickel-Ionen mit einem $[Fe_4S_4]$-Cluster verbrückt sind. Die Biosynthese von komplexen Metallclustern wird normalerweise durch mehrere akzessorische Proteine katalysiert. Für ACS war bisher jedoch nicht bekannt, welche Proteine an der Maturation des A-Clusters beteiligt sind.

In der vorliegenden Arbeit wurde das Protein AcsF als Maturationsfaktor der ACS identifiziert und sowohl biochemisch als auch strukturell charakterisiert. Es konnte gezeigt werden, dass apoACS und AcsF aus *Carboxydothermus hydrogenoformans* einen stabilen Komplex mit einer Stöchiometrie von 1:2 bilden. Der Komplex ist in der Lage zwei Nickel-Ionen zu binden, wohingegen apoACS unter den gleichen Bedingungen im Schnitt nur weniger als ein Nickel-Ion bindet. Durch die Komplexbildung wird also entweder eine neue Nickelbindungsstelle geschaffen oder die Affinität der vorhandenen Bindungsstellen in apoACS erhöht. Obwohl zwei Nickel-Ionen an den ACS-AcsF-Komplex gebunden sind, zeigt dieser keine Acetyl-CoA Synthase-Aktivität. Die inaktive Spezies kann durch Zugabe von Mg-ATP in eine aktive Form überführt werden. Zusammengenommen zeigen die Ergebnisse, dass AcsF die Mg-ATP abhängige Nickel-Insertion in ACS katalysiert.

Weiterhin wurden sequenzielle und strukturelle Unterschiede zwischen AcsF-Proteinen und CooC-Proteinen, den Maturationsfaktoren der Kohlenmonoxid Dehydrogenase,

aufgedeckt. AcsF- und CooC-Proteine gehören laut InterPro Datenbank zur gleichen Familie (IPRO14433). Ein Sequenzähnlichkeits-Netzwerk konnte zeigen, dass diese in drei Untergruppen unterteilt werden kann, wobei AcsF- und CooC-Proteine jeweils eine eigene Untergruppe bilden. Die Kristallstrukturen von AcsF aus *C. hydrogenoformans* und AcsF aus *Archaeoglobus fulgidus* wurden gelöst und mit zwei Strukturen von CooC-Proteinen verglichen. Die größten strukturellen Unterschiede zwischen den AcsF-und CooC-Proteinen sind zwischen dem Switch I Motif und dem CXC Motif zu erkennen.

Abstract

Acetyl-CoA synthase (ACS) is a key enzyme in the reductive acetyl-CoA pathway that catalyzes the reaction of CO, CoA and a methyl cation to form acetyl-CoA. The reaction takes place at a Ni,Ni-[Fe_4S_4] cluster (A-cluster), in which two nickel ions are bridged to a [Fe_4S_4] cluster. The biosynthesis of complex metal clusters is usually catalyzed by several accessory proteins. For ACS, however, so far it was not known which proteins are involved in the maturation of the A-cluster.

In this work, the protein AcsF was identified as maturation factor of ACS, and it was characterized both biochemically and structurally. It was shown that apoACS and AcsF from *Carboxydothermus hydrogenoformans* form a stable complex with a stoichiometry of 1:2. The complex is able to bind two nickel ions, whereas apoACS binds on average only less than one nickel ion. Even though two nickel ions are bound to the ACS-AcsF complex, the complex does not have acetyl-CoA formation activity. The inactive species can be converted to an active form by addition of Mg-ATP. Taken together, these results show that AcsF catalyzes the Mg-ATP dependent insertion of nickel into ACS.

Furthermore, differences in structure and sequence between AcsF proteins and CooC proteins were uncovered. CooC proteins are maturation factors of carbon monoxide dehydrogenase and belong to the same family as AcsF proteins as classified by the InterPro database (IPRO14433). A sequence similarity network showed that the family can be divided into three subgroups, of which both CooC and AcsF proteins form their own subgroup. The crystal structures from AcsF from *C. hydrogenoformans* and AcsF from *Archaeoglobus fulgidus* were solved and compared with two structures of CooC

proteins. The comparison revealed that the main structural differences between AcsF and CooC proteins are present between the switch I motif and the CXC motif.

Contents

Zusammenfassung **iii**

Abstract **v**

1 Introduction **1**

1.1 Nickel containing enzymes 1

1.2 Acetyl-CoA synthase – a key player in the reductive acetyl-CoA pathway 5

1.2.1 The reductive acetyl-CoA pathway 5

1.2.2 Acetyl-CoA synthase 7

1.3 Maturation of nickel containing enzymes 11

1.3.1 Maturation of [NiFe]-hydrogenase 11

1.3.2 Maturation of urease 13

1.3.3 Maturation of CODH 14

1.4 The SIMIBI class of proteins 14

1.4.1 HypB proteins 15

1.4.2 UreG proteins 17

1.4.3 CooC proteins 17

1.5 Research objectives 18

2 Materials and Methods **19**

2.1 Chemicals and enzymes 19

2.2 Other materials 19

2.3 Anoxic work 19

2.4 Microbiology 20

2.4.1 Media and antibiotics 20
2.4.2 Bacterial strains 21
2.5 Molecular biology 21
2.5.1 Plasmid transformation 21
2.5.2 Molecular cloning 21
2.5.3 Agarose gel electrophoresis 24
2.5.4 Isolation of plasmid DNA 24
2.6 Heterologous expression of proteins 24
2.6.1 Expression of genes encoding for AcsF and CooC proteins 24
2.6.2 Expression of the gene encoding for ACS 25
2.6.3 Co-expression of the genes encoding for ACS and $AcsF_{Ch}$ 25
2.7 Protein purification 26
2.7.1 Purification of $AcsF_{Ch}$ 26
2.7.2 Purification of $AcsF_{Af}$ 26
2.7.3 Purification of $CooC3_{Ch}$ 27
2.7.4 Purification of ACS 27
2.7.5 Sodium dodecyl sulfate polyacrylamide gel electrophoresis 27
2.8 UV-Vis spectroscopy 28
2.9 Acetyl-CoA formation activity of ACS 28
2.9.1 Synthesis of methylcobinamide 28
2.9.2 Reconstitution of apoACS 28
2.9.3 Measurements of acetyl-CoA formation activity 28
2.10 Measurements of NTPase activity 29
2.10.1 Coupled enzyme assay 29
2.10.2 Malachite green assay 29
2.11 Oligomerization and complex formation behaviour 30
2.11.1 Calibration of the size exclusion column 30

2.11.2 Analysis of the dimerization behaviour of AcsF$_{Ch}$ in the presence of nucleotide . . . 30
2.11.3 Analysis of the complex formation behaviour of ACS with AcsF$_{Ch}$, K10A-AcsF$_{Ch}$, C108A-AcsF$_{Ch}$, CooC1$_{Ch}$ and CooC3$_{Ch}$ . . . 30
2.12 Metal binding analysis . . . 31
2.12.1 UV-Vis spectroscopic nickel binding assay . . . 31
2.12.2 Isothermal titration calorimetry . . . 31
2.12.3 Removal of excess metal after ITC experiments . . . 32
2.12.4 Metal analysis by inductively coupled plasma optical emission spectroscopy . . . 32
2.13 Sequence similarity analysis and alignment visualization . . . 32
2.14 Protein crystallization and structure determination . . . 33
2.14.1 Data collection and processing . . . 33
2.14.2 Crystallization and data collection of CooC3$_{Ch}$ . . . 33
2.14.3 Structure determination . . . 34
2.14.4 Structure representation . . . 34

3 Results and discussion **35**
3.1 Protein expression and purification . . . 35
3.1.1 Expression and purification of AcsF$_{Ch}$ . . . 35
3.1.2 Expression and purification of AcsF$_{Af}$ . . . 36
3.1.3 Expression and purification of ACS . . . 37
3.2 UV-Vis spectral features of AcsF$_{Ch}$ and ACS . . . 39
3.3 Nickel binding behaviour of AcsF$_{Ch}$ . . . 39
3.3.1 Spectroscopic nickel binding assay . . . 39
3.3.2 Isothermal titration colorimetry of AcsF$_{Ch}$ with nickel . . . 41
3.4 Dimerization behaviour of AcsF$_{Ch}$ in the presence of nucleotides . . . 41
3.5 ATPase activity of AcsF$_{Ch}$ . . . 42
3.6 Complex formation behaviour of ACS . . . 44

3.6.1 Complex formation of apoACS with $AcsF_{Ch}$ 44
3.6.2 Complex formation of ACS with apo$AcsF_{Ch}$ variants and CooC proteins 44
3.7 Reconstitution of ACS 46
3.7.1 Reconstitution of apoACS by $AcsF_{Ch}$, Mg-ATP and $NiCl_2$ 46
3.7.2 Reconstitution of ACS by apo$AcsF_{Ch}$ variants and CooC proteins . 48
3.8 Nickel binding behaviour of apoACS and the apoACS-$AcsF_{Ch}$ complex . 49
3.8.1 Nickel titration by ITC 49
3.8.2 Metal analysis of Ni-titrated samples by ICP-OES 49
3.9 Conversion of the inactive Ni-ACS-$AcsF_{Ch}$ complex into active ACS . . . 51
3.10 Detailed sequence analysis of AcsF- and CooC-type proteins 51
3.11 Protein crystallization and structure determination 55
3.11.1 Crystallization and structure determination of $AcsF_{Ch}$ 55
3.11.2 Crystallization and structure determination of $AcsF_{Af}$ 56
3.11.3 Structure determination of $CooC3_{Ch}$ 57
3.12 Crystal structures 58
3.12.1 Overall structure of $AcsF_{Ch}$ 58
3.12.2 Overall structure of $AcsF_{Af}$ 61
3.12.3 Overall structure of $CooC3_{Ch}$ 63
3.12.4 Nickel binding properties of $CooC3_{Ch}$ 63
3.13 Structural comparison of AcsF and CooC proteins 65

4 Conclusion **71**

A Appendices **75**
A.1 Appendices to Materials and Methods 75
A.2 Appendices to Results and Discussion 78

Bibliography **93**

List of Figures **109**

List of Tables **112**

List of Abbreviations **113**

Declaration **117**

Acknowledgement **119**

1 | Introduction

1.1 Nickel containing enzymes

While the biological relevance of some metals, such as iron, zinc and copper, was established by the early 20th century [1, 2, 3, 4], it was only in the 1960s that the significance of nickel was recognized. Bartha et al. showed that some bacteria of the genus *Alcaligenes* required nickel to grow chemolitotrophically [5], while Bertrand et al. identified nickel as a growth stimulator for the alga *Chlorella vulgaris* [6]. At that time, however, the exact function of nickel remained unknown. It took another decade until it was shown that nickel is an integral part of certain enzymes and crucial to their function.

Today, eleven enzymes are known to function with nickel. These are acetyl-CoA-synthase (ACS) [7], acireductone dioxygenase [8], carbon monoxide dehydrogenase (CODH) [9], glycerol-1-phosphate dehydrogenase in *Bacillus subtilis* [10], glyoxalase I in *Escherichia coli* [11] or *Leishmania major* [12], lactate racemase [13], methyl-CoM reductase [14], nickel superoxide dismutase [15], [NiFe]-hydrogenase [16], urease [17], and quercetinase in *Streptomyces* sp. FLA [18].

From both a structural and functional point of view, nickel-containing enzymes are highly diverse (Fig. 1.1). In acireductone dioxygenase, glyoxalase I and quercetinase [19, 20, 21], for example, nickel is coordinated in a simple octahedral manner. In ACS, CODH and [NiFe]-hydrogenase, on the other hand, nickel is part of a unique metal cluster [7, 22, 23]. In ACS, two nickel ions are bridged to a $[Fe_4S_4]$ cluster. One of the nickel ions is coordinated by two backbone nitrogen atoms and two bridging cysteine

residues. The other nickel ion is also coordinated by the two bridging cysteines, plus a third cysteine residue and a water molecule. Both nickel ions are coordinated in a square planar fashion. In CODH the nickel ion is bound to a [3Fe4S] subcluster. It is coordinated in a T-shape to a cysteine thiolate and two inorganic sulfides. In hydrogenase, the nickel ion is part of a $NiFe(CN)_2CO$ cluster. The nickel ion is coordinated by a bridging oxygenic species, two bridging cysteine residues and two cysteine thiolates.

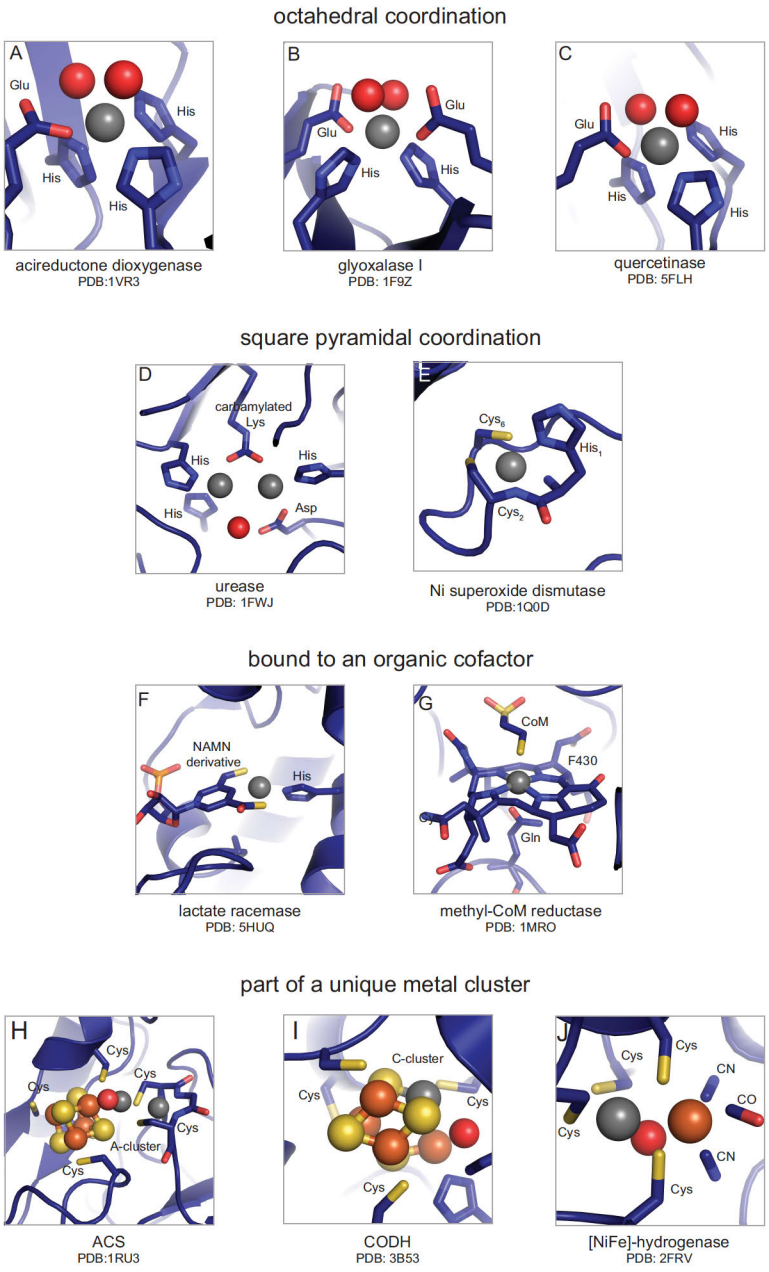


Figure 1.1: Overview of the active sites of currently known nickel-containing enzymes. Nickel and iron are depicted as grey and orange spheres, respectively. Oxygen atoms of water molecules are shown as red spheres. Nickel adopts an octahedral conformation in the enzymes acireductone dioxygenase (A), glyoxalase I (B) and quercetinase (C). Square pyramidal coordinations of nickel are found in urease (D) and nickel superoxide dismutase (E). In lactate racemase (F) and methyl-CoM reductase (G) the nickel ion is coordinated to an organic cofactor. In ACS (H), CODH (I) and [NiFe]-hydrogenase (J), the nickel ion is part of a unique metal cluster.

Table 1.1: Nickel containing enzymes and their corresponding reactions.

	involved in	**reaction**
ACS	carbon metabolism	$CH_3 + CO + CoA \rightleftharpoons$ acetyl-CoA
acireductone dioxygenase	methionine metabolism	1,2-dihydroxy-3-keto-5-methylthiopentene + $O_2 \rightleftharpoons$ 3-methylthiopropanoate + $HCOO^-$ + CO
CODH	carbon metabolism	$CO_2 + 2\ e^- + 2\ H^+ \rightleftharpoons CO + H_2O$
glycerol-1-phosphate dehydrogenase	biosynthesis of phospholipids	dihdroxyacetone phosphate $\rightleftharpoons$ *sn*-glycerol-1-phosphate
glyoxalase I	detoxification of methylglyoxal	methylglyoxal + gluthathione $\rightleftharpoons$ *S*-D-lactoyl-glutathione
lactate racemase	pyruvate metabolism	L-lactic acid $\rightleftharpoons$ D-lactic acid
methyl-CoM reductase	methanogenesis	Me-S-CoM + CoB-SH $\rightleftharpoons CH_4$ + CoB-S-S-CoMe
nickel superoxide dismutase	oxidative stress response	$2\ O_2^- + 2\ H^+ \rightleftharpoons H_2O_2 + O_2$
[NiFe]-hydrogenase	energy metabolism	$H_2 \rightleftharpoons 2\ H^+ + 2\ e^-$
urease	nitrogen metabolism	$(NH_2)_2CO + H_2O \rightleftharpoons CO_2 + 2\ NH_3$
quercetinase	degradation of flavonoids	quercetin + $O_2 \rightleftharpoons$ 2-protocatechuoylphoroglucinol carboxylic acid + CO

Nickel-containing enzymes catalyze a variety of reactions, including condensation reactions, racemezations and redox reactions. Although nickel-containing enzymes are relatively few, they play an important role in the earth's biogeochemical cycles. Urease, for example, takes part in the nitrogen cycle, while ACS and CODH catalyze reactions of the global carbon cycle. A complete list of reactions catalyzed by nickel-containing enzymes is shown in Table 1.1.

Nickel-containing enzymes are highly abundant in microorganisms. A genomic analysis by Zhang et al. revealed that 59 % of 540 bacterial and 82 % of 47 archeal genomes contain enzymes or transporters associated with nickel [24]. By contrast, Ni-containing enzymes are less prevalent in eukaryotes: Apart from glyoxalase I in *Leishmania major*, the only Ni-containing enzyme found in eukaryotes is urease, which is present mainly in fungi, and also in some plants. Animals, according to the current state of knowledge, do not possess any Ni-containing enzymes.

1.2 Acetyl-CoA synthase – a key player in the reductive acetyl-CoA pathway

1.2.1 The reductive acetyl-CoA pathway

The reductive acetyl-CoA pathway is one of the six autotrophic carbon fixation pathways [25]. Two molecules of CO_2 enter this pathway. One is reduced to CO by CODH, while the other is reduced to a cofactor-bound CH_3-moiety (Fig. 1.2). ACS catalyzes the final and most critical reaction of this pathway: the formation of acetyl-CoA from CO, the CH_3-cation and CoA . This pathway is often referred to as the Wood-Ljungdahl pathway, after the researchers who did pioneering work in this field.

Of the six pathways, the reductive acetyl-CoA pathway is the most energetically efficient in carbon fixation (Table 1.2). During the many steps toward the synthesis of a pyruvate molecule, only one molecule of ATP is hydrolyzed. However, the enzymes operating in this pathway are highly sensitive to oxygen. For this reason, organisms utilizing this pathway – carboxydotrophic and acetogenic bacteria, methanogenic archaea, and sulfate-reducing bacteria and archaea – are confined to ecological niches and live under extreme conditions in the absence of dioxygen. The carboxydotrophic bacterium *Carboxydothermus hydrogenoformans*, for example, thrives at temperatures around 70 °C and high concentrations of CO (up to 100 % v/v in the gasphase) [26]. The sulfate-reducing archaeon *Archaeoglobus fulgidus* prefers even higher growth temperatures (76 °C – 83 °C). It can be found in marine hydrothermal systems, and also in the production fluids of oil reservoirs [27, 28].

The reductive acetyl-CoA pathway is used not only for carbon fixation from C1-compounds; in some organisms, it also operates in the reverse direction. *A. fulgidus*, for example, can grow lithoautotrophically on H_2, thiosulfate and CO_2 [29], but prefers to grow heterotrophically on lactate and sulfate [30]. In the latter case, the pathway is operated in the reverse direction. Pyruvate is oxidized to CO_2 and acetyl-CoA, which is then further degraded, via the reductive acetyl-CoA pathway, to CoA and two molecules of CO_2.

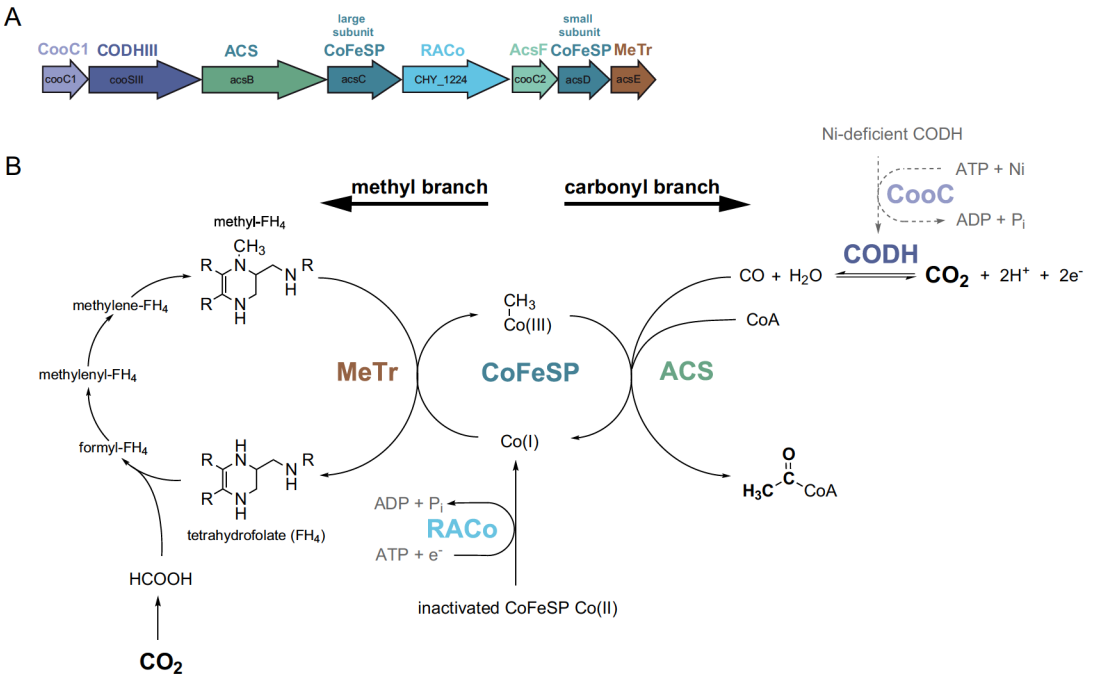


Figure 1.2: The reductive acetyl-CoA pathway in *C. hydrogenoformans*. A) Gene cluster of the reductive acetyl-CoA pathway in *C. hydrogenoformans*, which contains the genes encoding for CODH, ACS, the corrinoid iron sulfur protein (CoFeSP) and methyltransferase (MeTr) as well as the accessory genes encoding for CooC1, RACo (reductive activator of CoFeSP) and AcsF. Genes encoding for enzymes that are part of the methyl-branch are located further downstream and are not shown. A similar figure was used in [31]. B) Overall reaction scheme of the reductive acetyl-CoA pathway including the accessory proteins RACo and CooC. The pathway can be divided into the methyl-branch and the carbonyl-branch. The methyl-branch differs between bacteria and archaea with regard to the enzymes and coenzymes used. Bacteria use tetrahydrofuran as the methylcarrier, while archaea use tetrahydropterin. In bacteria CO_2 is reduced to formate by formate dehydrogenase, while in archaea CO_2 is reduced to formyl-methanofuran by formylmethanofuran dehydrogenase.

Another function of this pathway is energy conservation. Acetogens, for example, can convert acetyl-CoA to acetyl-phosphate, which is then further converted to acetate, yielding one ATP molecule [32]. Methanogens can extract methyltetrahydromethanopterin from the pathway and reduce the methyl group to methane. This reaction

Table 1.2: Autotrophic carbon fixation pathways. The table was adapted from [25].

	ATP	reductant	found in
reductive pentose phosphate cycle	7	NAD(P)H	plants, algae, cyanobacteria, some proteobacteria
reductive citric acid cycle	2	NAD(P)H, ferredoxin	bacterial phyla Chlorobi, Aquificae and Nitrospirae, some proteobacteria
dicarboxylate-4-hydroxybutyrate cycle	5	NAD(P)H, ferredoxin	archaeal orders Desulfurococcales and Thermoproteales
3-hydroxypropionate-4-hydroxybutyrate cycle	9	NAD(P)H	archaeal order Sulfolobales, maybe archaea from the phylum Thaumarchaeota
3-hydroxypropionate bicycle	7	NAD(P)H	bacterial family Chloroflexaceae
reductive acetyl-CoA pathway	max 1	NAD(P)H, ferredoxin, $F420H_2$	archaeal phylum Euryarchaeota, bacterial phyla Firmicutes, Planctomycetes and Proteobacteria

is coupled to the formation of a transmembrane gradient, which is thought to yield 1.5 molecules of ATP per molecule of methane [33].

1.2.2 Acetyl-CoA synthase

ACS is the central enzyme of the reductive acetyl-CoA pathway. It catalyzes the reaction of CO, CoA and a CH_3-cation to form acetyl-CoA. The active site is a Ni,Ni-$[Fe_4S_4]$ cluster, which is also called the A-cluster [7]. ACS can be present either as the monofunctional enzyme or in a complex with other enzymes from the acetyl-CoA pathway. From bacterial sources ACS has been isolated both as a monomer and in a bifunctional complex with CODH [7, 34]. In archaea, ACS is part of the large multienzyme complex, acetyl-CoA decarbonylase synthase (ACDS), which is believed to be present in an $(\alpha_2\varepsilon_2)_4(\beta)_8(\gamma\delta)_8$ arrangement, where the $\alpha_2\varepsilon_2$-subunits have CODH functionality, the β-subunit corresponds to ACS and the $\gamma\delta$-subunits have CoFeSP and methyltransferase functionality [35, 36, 37].

Monofunctional ACS consists of three domains: the N-terminal domain, the middle domain and the C-terminal domain (Fig. 1.3 A). The A-cluster is coordinated exclusively by residues from the C-terminal domain. The functional enzyme contains two nickel ions in a tetrahedral coordination geometry [7]. These are bridged to a [Fe_4S_4] cluster (Fig. 1.3 B). The [Fe_4S_4] cluster is a typical cubane-type cluster coordinated by four cysteine residues. The nickel ion located distal to the [Fe_4S_4] cluster, denoted Ni_d, is coordinated by two nitrogen atoms from the protein backbone and two thiol ligands of a Cys-Gly-Cys motif. The other nickel ion, denoted Ni_p due to its location proximal to the [Fe_4S_4] cluster, is coordinated by three cysteine residues and a forth ligand. Ni_p is labile, i.e. it can be removed by chelating agents and replaced by other metals [38].

Bifunctional ACS/CODH is a heterotetramer. Two CODH subunits are flanked on each side by an ACS subunit (Fig. 1.3 C). The active sites of CODH and ACS are connected by a tunnel network, which ensures that CO generated at the active site of CODH is directly transferred to the A-cluster of ACS [39]. Two crystal structures of bifunctional ACS/CODH have been solved. The structure from Doukov et al. contained two Ni,Cu-[Fe_4S_4] clusters [34], while the structure from Darnault et al. contained a Ni,Ni-[Fe_4S_4] cluster in one subunit and a Ni,Zn-[Fe_4S_4] cluster in the other [39]. In these structures, zinc and copper occupy the proximal positions and adopt a distorted tetrahedral coordination. When these crystal structures were published, the composition of the active site was still debated. It has since become clear that copper- and zinc-containing forms of ACS are inactive, and that only ACS with a Ni,Ni-[Fe_4S_4] cluster is functional. In the structure from Darnault et al., ACS adopts two different conformations, open and closed, which differ by a rotation of the N-terminal domain of approximately 50° relative to the middle and C-terminal domain (Fig. 1.3 D). In the open conformation, the tunnel from CODH to the A-cluster is blocked, and the A-cluster is more solvent exposed. The open conformation resembles the structure of monomeric ACS from *C. hydrogenoformans* containing the Ni,Ni-[Fe_4S_4] cluster, while the closed conformation is similar to the structure from Doukov et al. containing the Ni,Cu-[Fe_4S_4] cluster.

The structure of the archaeal counterpart, the β-subunit of the ACDS complex, is not known. Sequence comparison of the archaeal β-subunit with bacterial ACS suggests that the β-subunit resembles the middle and C-terminal domain of bacterial ACS, while the N-terminal domain is lacking [40].

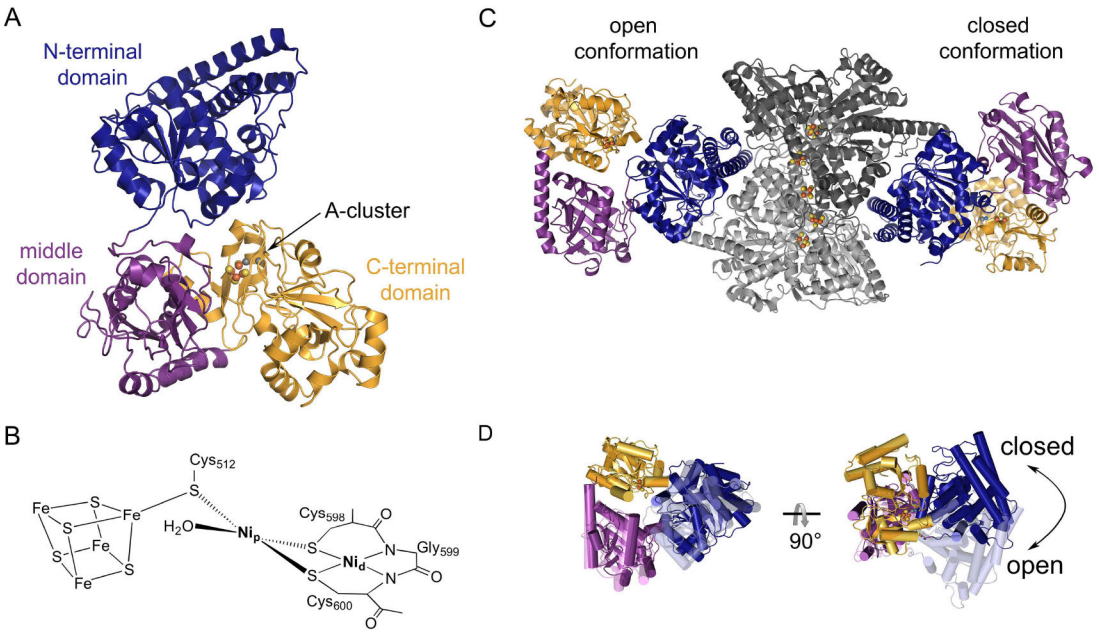


Figure 1.3: Crystal structures of ACS. A) Monomeric ACS from *C. hydrogenoformans* (PDB: 1R3U). The N-terminal domain is shown in blue, the middle domain in violet, and the C-terminal domain in yellow. The two nickel ions are depicted as grey spheres. The N-terminal domain contains a Rossmann fold, whereas the middle and C-terminal domain do not show any similarities with known structural motifs. B) Schematic representation of the A-cluster of ACS. C) Bifunctional ACS/CODH from *M. thermoacetica* (PDB: 1OAO). The subunits of CODH are shown in dark and light grey. ACS has the same coloring scheme as in A). The zinc ion is shown as blue sphere in the closed conformation. D) Overlay of ACS conformations of the open and closed state, which are shown as transparent and opaque cartoon representations, respectively.

From a chemical point of view, the condensation of three substrates to acetyl-CoA is a very challenging reaction. Although many studies have attempted to elucidate the ACS reaction mechanism, the details remain unclear. It is generally agreed that

the substrates bind to Ni_p, that CoA is the last substrate to bind, and that the $[Fe_4S_4]$ cluster and Ni_d remain in the +2 oxidation state during the reaction cycle. Yet to be determined, however, are the oxidation state of Ni_p during catalysis and the binding order of CH_3^+ and CO. One model, known as the paramagnetic mechanism, includes the paramagnetic Ni^{1+}-CO species, which has been detected by EPR, as a catalytically active species in the reaction mechanism (Fig. 1.4) [41, 42]. It suggests a random order of binding of the CH_3^+ and CO and implies that one electron is shuttled from a short lived Ni^{3+} species to Ni^{2+} by a ferredoxin. The other model, the diamagnetic mechanism, assumes that the Ni^{1+}-CO species is an inhibited state and that CH_3^+ is the first substrate to bind [43, 40]. In this model the oxidation state of nickel cycles between 0 and +2 .

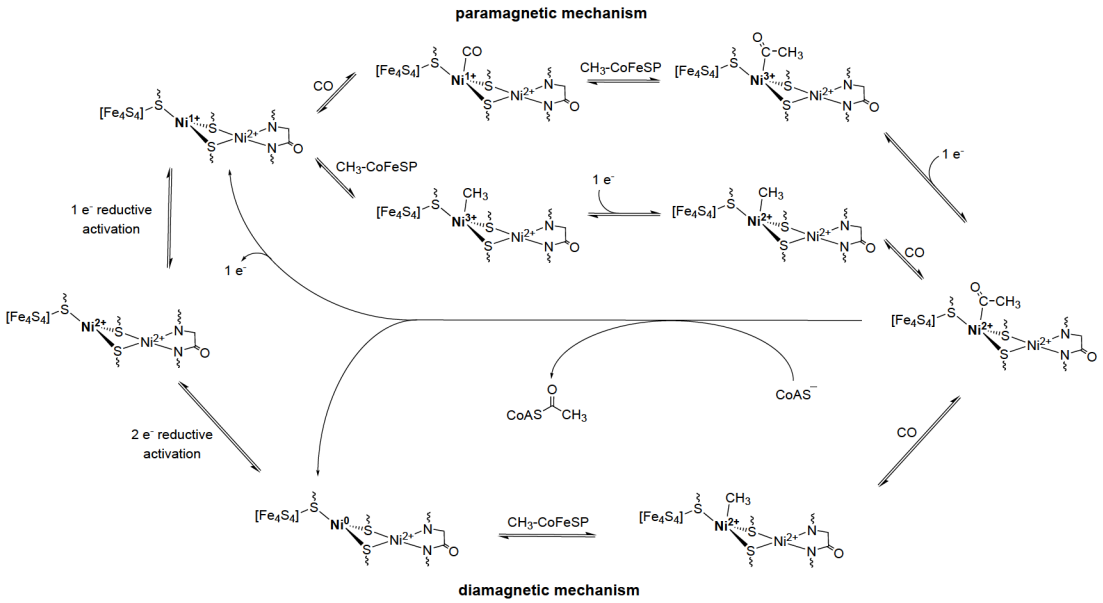


Figure 1.4: Two proposed mechanisms of ACS.

The reaction catalyzed by ACS is often considered the biological equivalent of the Monsanto or Cativa processes, which are the two main industrial processes for the synthesis of acetic acid [44, 45]. These processes are based on rhodium and iridium catalysts, respectively, which catalyze the carbonylation of methanol to acetic acid. Rhodium and iridium are both very rare and expensive metals. In biological systems the reaction is catalyzed by nickel and iron, metals that are readily available. This makes

ACS an interesting candidate for future biotechnological applications. Limiting factors, however, are its sensitivity to oxygen and the current lack of heterologous expression systems which produce active ACS. This lack is partly due to a poor understanding of the maturation of the A-cluster of ACS.

1.3 Maturation of nickel containing enzymes

Nickel-containing enzymes often require several accessory proteins. These proteins are involved in metal processing and the correct assembly of the active site. While the maturation of some nickel-containing enzymes, such as ACS, has not been investigated at all, the maturation systems of other enzymes, such as [NiFe]-hydrogenase, have been studied for nearly two decades. In the case of at least four enzymes, namely [NiFe]-hydrogenase, urease, lactate racemase and CODH, one of the maturation proteins is a GTPase or an ATPase. Many studies have attempted to elucidate the intricate interplay between maturation proteins within a particular system. The results of these studies are outlined in the following paragraphs.

1.3.1 Maturation of [NiFe]-hydrogenase

The maturation system of [NiFe]-hydrogenase is particularly complicated: it involves six proteins, which are encoded by the *hyp*ABCDEF operon (*hyp* for operon, which affects all three [NiFe]-hydrogenases in a pleiotropic manner) [46]. The maturation takes place in three steps: 1) assembly and insertion of $Fe(CN)_2CO$ into the large subunit, 2) nickel insertion, and 3) proteolytic cleavage of the large subunit followed by complex formation with the small subunit (Fig. 1.5) [47]. The first step, the assembly and insertion of the $Fe(CN)_2CO$ cofactor, requires four of the six accessory proteins, HypC, HypD, HypE and HypF [47]. The second step, nickel insertion, is catalyzed by HypA, HypB and, at least in some organisms, SlyD [48]. In *Helicobacter pylori* and *E. coli*, SlyD functions as a nickel storage protein, which can transfer a nickel ion to HypB [49, 50]. HypB is a GTPase and a member of the G3E family of the SIMIBI class of NTPases (see

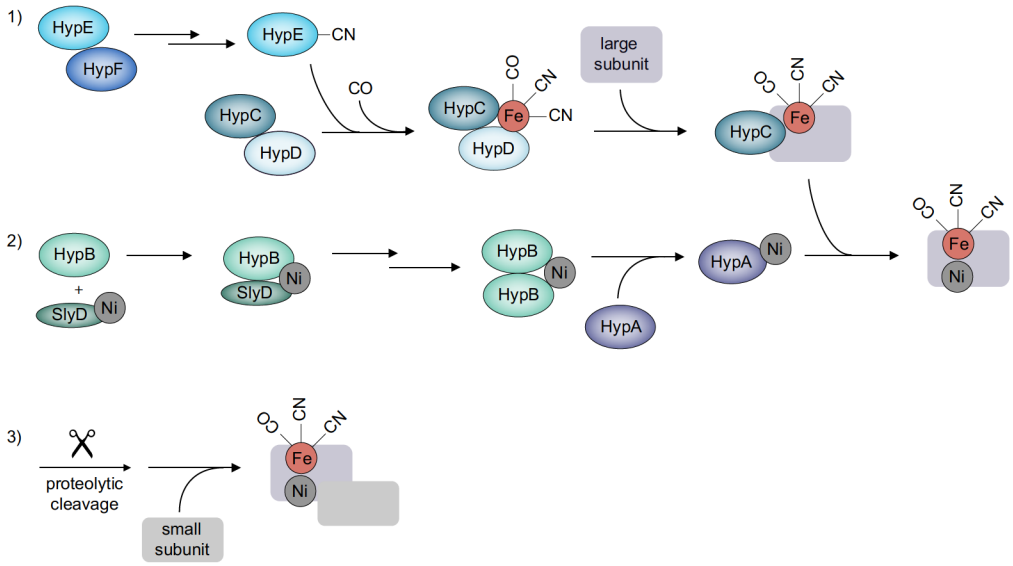


Figure 1.5: Proposed pathway of the maturation of [NiFe]-hydrogenase. After assembly of $Fe(CN)_2CO$ by the proteins HypC, HypD, HypE and HypF (1), nickel is inserted into hydrogenase (2). Nickel can be transferred from SlyD to HypB to HypA. Proteolytic cleavage of the large subunit followed by complex formation with the small subunit yields active [NiFe]-hydrogenase (3).

section 1.4). Early on, it was shown that the GTPase activity of HypB is crucial to the successful maturation of [NiFe]-hydrogenase [51]. HypB can interact with HypA. It can also transfer a nickel ion to HypA [52]. How exactly these processes are regulated by GTP hydrolysis is not yet clear. HypA is thought to be the protein which catalyzes the final step of nickel delivery into the hydrogenase precursor. This would agree with the finding that HypA forms a complex with the large subunit of hydrogenase [53]. The exact details of how nickel is transferred from HypA to the hydrogenase precursor are not yet clear. In the final step, the C-terminal of the large subunit is cleaved by a specific endopeptidase [54].

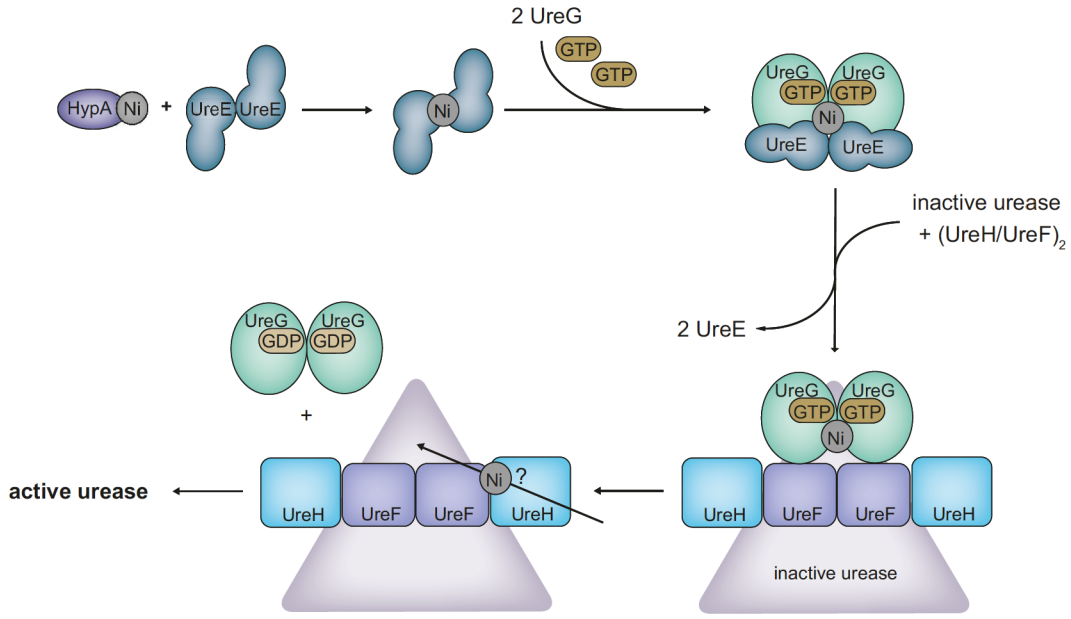


Figure 1.6: Proposed pathway of nickel delivery into *H. pylori* urease.

1.3.2 Maturation of urease

The delivery of nickel into urease involves at least four accessory proteins: UreE, UreF, UreG and UreH. Furthermore, the accessory proteins HypA and HypB, which are typically associated with hydrogenase maturation, seem to be involved in the maturation of urease (Fig. 1.6). Benoit et al. showed that HypA interacts with, and is able to transfer a nickel ion to, dimeric UreE [55]. In the presence of magnesium and GTP, dimeric UreE can form a complex with UreG, thereby transferring the nickel ion to UreG [56]. UreG, like HypB, is a GTPase of the G3E family of the SIMIBI class of NTPases (see section 1.4) [57]. Nickel-loaded UreG is thought to bind to the UreF-UreH-urease complex, which consists of a dimer of heterodimers [58]. Upon GTP hydrolysis, which can be stimulated by bicarbonate, nickel is released from UreG. GTP hydrolysis also leads to the dissociation of UreG from the UreF-UreH complex [59]. The final steps of the maturation pathway have not been determined in detail, however it was recently suggested that nickel may be channelled into the active site of urease through a tunnel network [60].

1.3.3 Maturation of CODH

Early studies in *Rhodospirillium rubrum* identified three proteins, CooC, CooJ and CooT, as maturation factors for CODH [61, 62]. CooJ is a histidine-rich protein that can bind up to four nickel ions. Together with CooC, it has been associated with nickel delivery [63]. CooT is only 7.1 kDa in size and its precise function is unclear. Homologous proteins of CooJ and CooT are found almost exclusively in purple non-sulfur bacteria of the family Rhodospirillaceae. They are absent in most other organisms containing a CODH. CooC proteins, on the other hand, are widespread throughout bacteria and archaea, and have therefore been characterised in greater detail. It was shown that *R. rubrum* strains, which contain a mutation within the *cooC* gene, would only grow in a medium which contained unusually high amounts of nickel [62]. CooC is a nickel-binding ATPase of the MinD-type family of the SIMIBI class of proteins (see section 1.4) [64]. Its ability to hydrolyze ATP is essential to the generation of active CODH in vivo [65].

1.4 The SIMIBI class of proteins

All of the maturation systems described above contain one maturation protein of the SIMIBI class of GTPases. The SIMIBI class (short for Signal Recognition Particle, MinD and BioD) is one of the two classes of the GTPase superclass of proteins containing a P-loop fold [66]. Despite its name, the superclass encompasses not only GTPases, but also related ATPases. Proteins of the GTPase superclass contain a central β-sheet, which is surrounded by α-helices. They are characterised by at least three conserved motifs[1]: The P-loop or Walker A motif (also referred to as G1) – GXXXGK[ST] [67], the switch I motif (G2), which contains a conserved threonine or aspartate residue, and the switch II or Walker B motif (G3) – DXXG. GTPases also contain a guanine specificity loop [NT]KXD (G4), in which the aspartate residue forms hydrogen bonds with the N1 and N2 atoms of the guanine base [68]. ATPases do not have a common motif establishing base specificity. The main difference between the SIMIBI class and the second class, TRAFAC

[1] where X is any amino acid, and squared brackets stand for alternate possibilities

(short for Translation Factor), is the orientation of the strand next to the switch II region: it is parallel in the SIMIBI class and antiparallel in the TRAFAC class [66].

The switch I and switch II regions undergo conformational changes depending on whether NTP or NDP is bound to the protein. The regions are therefore central to the signal transduction mechanism of SIMIBI proteins [69]. For the SIMIBI NTPases involved in the maturation of metalloproteins, two questions arise: 1) Do they bind a metal ion? 2) If so, is the NTP hydrolysis cycle in some way relayed to the metal binding capabilities of the NTPase? The crystal structures of HypB, UreG and CooC have been determined and their metal binding properties investigated, giving us some insight into their mechanism of action.

1.4.1 HypB proteins

HypB belongs to the G3E family of the SIMIBI class, which is characterised by the replacement of the aspartate residue in the G3 motif with a glutamate residue. HypB proteins can be divided into three different subgroups depending on the number of metal binding sites. One metal binding site, which is located at the interface of two HypB dimers, is conserved in all species. Additional binding sites can be a histidine-rich stretch at the N-terminal, as found in *Bradyrhizobium japanicum* [70, 71], or a high affinity nickel binding site formed by N-terminal cysteine residues, as found in *E. coli*. For HypB, it has been shown that the GTP hydrolysis cycle and metal binding are closely intertwined: GTP hydrolysis is inhibited by nickel binding, and it has been shown that nickel is coordinated to a histidine residue in the GDP-bound state, but to a cysteine residue in the GTP-bound state. The crystal structures of HypB from *Methanocaldococcus janaschii* and *H. pylori* (Fig. 1.7) revealed that the switch II region, particularly Cys142, translates structural changes at the nucleotide binding site to the metal binding site, and vice versa [68, 72].

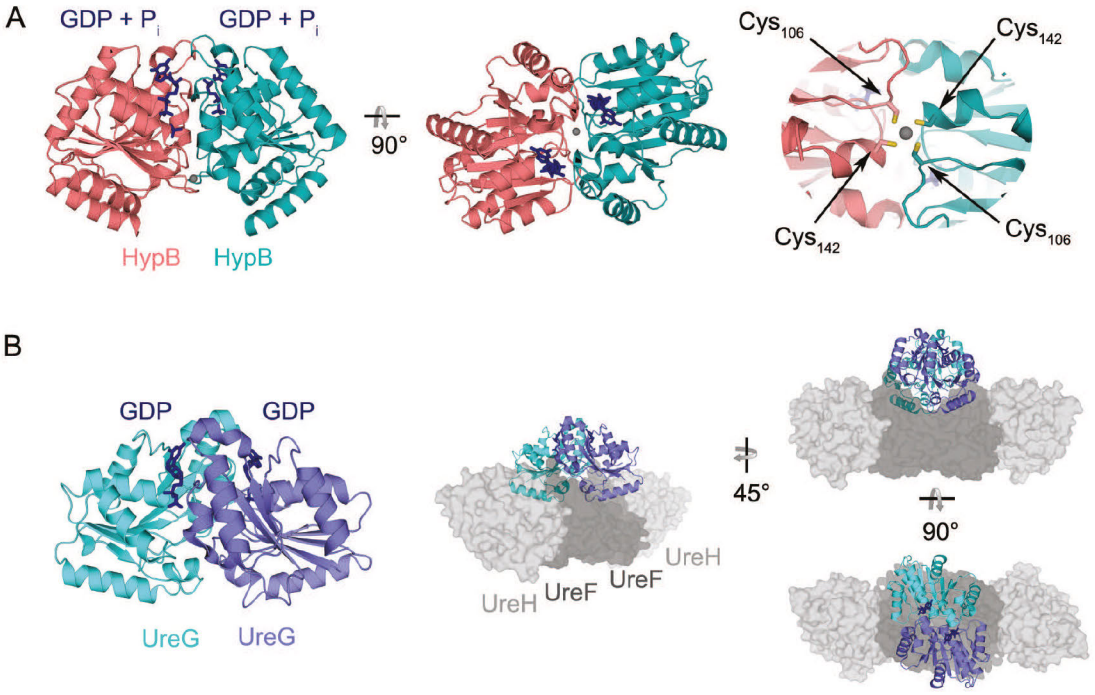


Figure 1.7: Crystal structures of HypB and UreG. A) Crystal structure of HypB from *H. pylori* (PDB: 4LPS). GDP and P_i are bound to the nucleotide binding site close to the dimer interface. The nickel ion is coordinated at the dimer interface in a square planar coordination geometry by four cysteine residues. B) Crystal structure of UreG from *H. pylori* in complex with UreF and UreH (PDB: 4HI0). As the structure of UreG alone is not available, the picture on the left shows a magnification of UreG in its conformation in the complex. A GDP molecule is bound to the nucleotide binding site. The UreG/UreF/UreH complex is also shown from three different perspectives, where UreF and UreH are coloured in dark and light grey, respectively.

As described, HypB is usually a GTPase of the G3E family. In some archaea of the order Thermococcales, however, a functional homologue of HypB, $HypB_{AT}$, which is a MinD-type ATPase, has been discovered [73]. $HypB_{AT}$ has no nickel binding site of its own, but upon ATP-dependent association of $HypB_{AT}$ with HypA, a new nickel binding site is formed in HypA. In this system the ATP hydrolysis cycle does not influence the metal binding properties of $HypB_{AT}$ itself, but rather influences the metal binding properties of another protein [74].

1.4.2 UreG proteins

UreG contains the highly conserved switch II motif (ESGGDNL) and is therefore also a member of the G3E family. While the structure of UreG was long unknown and its metal binding properties are poorly defined, recent progress has been made through the determination of the UreG/UreF/UreH complex structure and the accompanying biochemical studies [59]. These studies showed that nickel is released from UreG upon GTP hydrolysis, revealing a direct connection between metal and nucleotide binding. The crystal structure showed that the proposed metal binding site, consisting of a Cys-Pro-His motif, is positioned at the interface of the two UreG monomers, close to the interface of UreG and UreF. The switch I and II motifs are positioned between the GTP binding site and the metal binding site. This might explain how conformational changes that are triggered by GTP hydrolysis can affect the conformation of the metal binding site.

1.4.3 CooC proteins

CooC belongs to the MinD-type family of the SIMIBI proteins. MinD-type proteins are thought to be active only in the dimeric state. They contain a second lysine residue, referred to as the signature lysine, within the Walker A motif, which is located before the canonical lysine residue (G<u>K</u>GGXGK[T/S]). Crystal structures of hydrolysis deficient mutants of Soj and other members of the MinD family have shown that the signature lysine of one monomer interacts with the ATP molecule that is bound to the other monomer [75]. During ATP hydrolysis the signature lysine can balance the developing charges on the phosphate moiety, which explains why this residue is crucial to ATP hydrolysis and concomitant dimerization. This has also been confirmed for CooC, since a CooC variant in which the signature lysine was replaced by an alanine residue did not show ATPase activity and was unable to dimerize [64]. CooC proteins contain a conserved CXC motif, which forms a nickel binding site at the interface of two monomers. The exact coordination geometry of nickel has not yet been determined, but UV-Vis

spectroscopy experiments suggest that the coordination geometry of nickel differs in empty and nucleotide bound states of CooC. A mechanism was proposed based on the crystal structures and a derived model, which assumes that the metal binding site opens upon binding of ATP [76]. Although direct evidence remains elusive, there is a strong indication that the metal binding site is subjected to conformational changes during ATP hydrolysis.

1.5 Research objectives

Many biochemical and structural studies have been carried out on ACS, but its maturation system has not yet been investigated. It has been speculated that AcsF, a protein located in the same gene cluster as ACS, might be involved in the biosynthesis of the A-cluster of ACS [77]. So far, however, this is only conjecture. AcsF was once believed to be involved in the biosynthesis of the C-cluster in bifunctional ACS/CODH, but no evidence for this has been found [78]. The first objective of this thesis is to determine the function of AcsF and to establish whether AcsF acts as a maturation factor of ACS.

AcsF proteins belong to the same family as CooC proteins as defined by the InterPro database (family IPRO14433) [79].The second objective is to determine the sequential, structural and functional differences between AcsF proteins and CooC proteins.

2 | Materials and Methods

2.1 Chemicals and enzymes

All chemicals were at least of analytical grade and purchased either from AppliChem, Roche, Roth or Sigma-Aldrich. N_2, N_2/H_2 (95 %/5 %) and CO were purchased from Air Liquide. CO-saturated buffer was prepared by bubbling CO through the buffer for at least 10 min. Ti(III)citrate was prepared as previously described [80].

Restriction enzymes, DNA polymerases and T4 DNA ligase were purchased either from New England Biolabs (NEB) or Thermo Scientific. A pyruvate kinase and lactic dehydrogenase mixture, from rabbit muscle buffered in aqueous glycerol solution, was purchased from Sigma-Aldrich.

2.2 Other materials

Nucleic acid purification kits (plasmid purification, DNA gel extraction and PCR purification) were purchased from Thermo Scientific. Chromatography resins and columns were purchased from GE Healthcare or IBA. Bottles for metal-free buffers were rinsed with dilute HCl and then washed with Milli-Q H_2O. Chelex 100 resin (BioRad) was added to all metal-free buffers (5 to 10 g L^{-1}).

2.3 Anoxic work

All anoxic solutions were prepared at a Schlenk line in a bottle equipped with a butyl rubber septum by at least four cycles of evacuating and purging with N_2. Protein

purification, UV-Vis spectroscopy and crystallization experiments were carried out in anoxic glove boxes (model B; COY Laboratory Products) with an atmosphere of 5 % hydrogen and 95 % nitrogen. The glove boxes are equipped with palladium catalysts to remove traces of oxygen. ITC experiments were carried out in a glove box workstation (LABstar, MBRAUN) with an automatic inert gas purification system.

2.4 Microbiology

2.4.1 Media and antibiotics

Media for cultivation of *E. coli* had the composition shown in Table 2.1. LB medium and mTB medium were prepared manually. SOB medium was purchased premixed from Roth. Media were autoclaved for 20 min at 121 °C. For plasmid selection, appropriate antibiotics were added to the media with a working concentration of 50 µM mg^{-1} of kanamycin and carbenicillin, respectively.

Table 2.1: Composition of cultivation media.

name	composition in 1 L
lysogeny broth medium (LB)	10 g tryptone 5 g yeast extract 10 g NaCl
modified terrific broth medium (mTB)	12 g tryptone 24 g yeast extract 10 mL glycerol 10 % (v/v) phosphate buffer (0.17 M KH_2PO_4, 0.72 M K_2HPO_4)
super optimal broth medium (SOB)	20 g tryptone 5 g yeast extract 0.96 g $MgCl_2$ 0.5 g NaCl 0.186 g KCl
SOB medium with catabolic repression (SOC)	SOB with 20 mM glucose

2.4.2 Bacterial strains

For molecular cloning the *E. coli* strain DH5α (genotype: F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK-, mK+) *pho*A *sup*E44 λ-*thi*-1 *gyr*A96 *rel*A1) was used. Genes were expressed in *E. coli* BL21 (genoytpe: F– *omp*T *hsd*SB(rB-, mB-) *gal dcm* (DE3)). This strain is deficient in Ion and OmpT proteases and contains a T7 expression system. Neither *E. coli* DH5α nor *E. coli* BL21(DE3) has an intrinsic antibiotic resistance.

2.5 Molecular biology

2.5.1 Plasmid transformation

Chemically competent cells of either *E. coli* DH5α or *E. coli* BL21(DE3) were thawed on ice. Plasmid DNA was then added to the cell suspension, and the cells were kept on ice for another 15 min. Cells were subjected to a heat shock for 45 s at 42 °C and again incubated on ice for 2 min. 700 µL SOC medium was added to the cell suspension and the culture was incubated for 1 h at 37 °C. 100 µL cell suspension was plated on an LB agar plate containing the corresponding antibiotics. The plates were incubated for approximately 16 h at 37 °C.

2.5.2 Molecular cloning

For preparing plasmid DNA containing the gene of choice, either traditional cloning techniques using restriction enzymes or the polymerase incomplete primer extension (PIPE) method was used [81, 82]. For PIPE cloning, two polymerase chain reactions (PCRs) were set up. For the preparation of the vector, the PCR contained 0.1 ng template, 0.5 µM of the forward and reverse primer, and 25 µL Phusion High Fidelity PCR master mix (NEB) in a 50 µL reaction volume (primer sequences are listed in the Table A.1 in the Appendices). The DNA was amplified using the Mastercycler pro PCR machine (Eppendorf) with the protocol shown in Table 2.2. The PCR product was purified and aliquots were stored at -20 °C. Using appropriate primers and templates, the set up of the

PCR for the preparation of insert was the same as for the vector PCR. After purification of the insert, 1 µL of vector was annealed with 1 µL of insert for 1 min on ice and then directly transformed into *E. coli* DH5α.

Table 2.2: PCR protocol for PIPE cloning. The cycle consisting of denaturation, annealing and elongation was carried out 35 times.

	vector		insert	
	T [°C]	time [s]	T [°C]	time [s]
initial denaturation	98	120	98	120
denaturation	98	10	98	10
annealing	65	10	62	10
elongation	72	120	72	12
cooling	4		4	

Site-directed mutagenesis was carried out with the quikchange method [83]. The PCR contained 10 ng template, 0.5 µM of each primer, a dNTP mix and 2.5 U *Pfu* polymerase in the *Pfu* buffer provided from the supplier. The protocol was carried out as described in Table 2.3. After the PCR, the mixture was treated with 10 U Dpn1 for 1 h at 37 °C to digest the template DNA. 5 µL of the Dpn1-treated DNA was used for the transfrormation of *E. coli* DH5α.

Table 2.3: PCR protocol for quikchange mutagenesis. The cycle consisting of denaturation, annealing and elongation was carried out 18 times.

	T [°C]	time [s]
initial denaturation	95	60
denaturation	95	50
annealing	T_M	50
elongation	68	60
elongation	68	420

At the start of this project, four out of five genes were already present in cloning vectors. There was no need to clone these from genomic DNA. Nonetheless, during the course of the project, all four genes were cloned into different vectors, either to improve the expression of the gene or to allow a different purification strategy. The modified pET28a vectors used in this work, which contained either a streptag or a twin-streptag, were kindly provided by Dr. Martin Bommer.

Two plasmids containing the gene *cooC2* from *C. hydrogenoformans*, the product of which was named AcsF$_{Ch}$, were kindly provided by Dr. Jae-Hun Jeoung. The gene was present in the plasmid pET11a-AcsF as well as the plasmid pASK-IBA17-AcsF. In the former, the *cooC2* gene is located between the NdeI and BamHI restriction sites and the gene product has no tag; in the latter, the gene is located between the NarI and XhoI sites and the gene product will have a streptag. To improve expression, the gene was cloned into the modified pET28a vector with an N-terminal streptag and a TEV cleavage site (pET28aTEVstrep) by PIPE cloning. The new plasmid was named pET28a-AcsFstrep. The plasmid pET28a-AcsFstrep was later used for the generation of the two variants K10A-AcsF$_{Ch}$ and C108A-AcsF$_{Ch}$ by the quikchange method [84].

The *cooC1* gene from *A. fulgidus*, the product of which was named AcsF$_{Af}$, was ordered as a synthetic gene from Eurofins genomics. NdeI and BamHI restriction sites were added to the end of the genes as shown in Figure A.1 in the Appendices. The gene was provided in the vector pEX-A2. This vector was digested with NdeI and BamHI, and the gene was ligated into an NdeI/BamHI digested pET28aTEVstrep vector.

A plasmid containing the *cooC3* gene from *C. hydrogenoformans*, the product of which is CooC3$_{Ch}$, was kindly provided by Dr. Jae-Hun Jeoung. The gene was present in the plasmid pET20b-CooC3. The plasmid was digested with NdeI and HindIII and ligated into the pET28a-TEVstrep vector. The plasmid was named pET28a-CooC3strep.

A plasmid containing the *acsB* gene from *C. hydrogenoformans*, the product of which is ACS, was kindly donated by Dr. Sebastian Götzl. The *acsB* gene was present in the plasmid pET28a-ACSstrep. To improve the purification strategy, the gene was cloned into a pET28a-twinstrep vector, which contains a twin-streptag instead of a single streptag. The pET28a-ACSstrep plasmid was digested with NdeI and BamHI and the digestion product was ligated into a NdeI/BamHI digested pET28a-twinstrep vector. The plasmid was named pET28a-ACStwinstrep.

2.5.3 Agarose gel electrophoresis

PCR products or fragments from restriction digests were purified by agarose gel electrophoresis containing 1 % (w/v) agarose in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA pH 8.0). The gel was pre-stained with Roti® GelStain (Roth). DNA samples were mixed with an appropriate amount of 6x DNA loading dye (10 mM Tris-HCl pH 7.6, 0.03 % bromophenol blue, 0.03 % xylene cyanol FF, 60 % glycerol, 60 mM EDTA) and loaded on the gel. The GeneRuler 1 kb DNA ladder (Fermentas) was used as a size reference. The gel was run for 1 h at 100 V. The DNA fragments were excised from the gel and purified with the DNA gel extraction kit.

2.5.4 Isolation of plasmid DNA

A 5 mL culture of *E. coli* DH5α containing the desired plasmid was incubated for approximately 16 h at 37 °C. The culture was centrifuged and the plasmid DNA was isolated using the plasmid purification kit according to the manufacturer's protocol.

2.6 Heterologous expression of proteins

2.6.1 Expression of genes encoding for AcsF and CooC proteins

For expression of genes encoding for AcsF proteins, CooC proteins or variants thereof, the corresponding plasmids were transformed into *E. coli* BL21(DE3). Cells were grown aerobically in LB medium containing the appropriate antibiotics. After the optical

density at 600 nm (OD_{600}) reached 0.7 ± 0.1, the culture was induced with 0.2 mM isopropyl β-D-thiogalactopyranoside (IPTG) and the temperature was decreased to 28 °C. Cells were harvested 16 to 24 hours after induction and washed with 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM tris(2-carboxyethyl)phosphine (TCEP) and 2 mM EDTA. Cell pellets were frozen in liquid N_2 and stored at -80 °C.

2.6.2 Expression of the gene encoding for ACS

For expression of the gene encoding for ACS, *E. coli* BL21(DE3), containing the pET28a-ACSstrep or pET28a-ACStwinstrep plasmid, were cultivated in mTB medium containing the appropriate antibiotics. The culture was grown in 5 L bottles in a water bath at 37 °C. The culture was stirred aerobically and after one hour the medium was supplemented with filter sterilized 0.1 mM $FeSO_4$ and 0.5 mM cysteine. When the OD_{600} reached 0.6 ± 0.1, cultures were induced with 0.5 mM IPTG and transitioned to anaerobic growth by closing the bottles with a butyl rubber septum. After 24 hours cells were quickly harvested aerobically, frozen in liquid N_2 and stored at -80 °C.

2.6.3 Co-expression of the genes encoding for ACS and $AcsF_{Ch}$

A co-expression system was established in *E. coli* BL21(DE3) to produce both proteins, ACS and $AcsF_{Ch}$, simultaneously. To prevent plasmid loss from the system, two plasmids, each with its own antibiotic selection marker, were chosen. Cells containing the plasmid pET28a-ACSstrep with a kanamycin resistance marker were made competent by the method from Inoue et al. [85]. The competent cells were then transformed with the plasmid pET11a-AcsF, which contains an ampicillin resistance marker. The cells containing both plasmids were cultivated following the same protocol as described for the production of ACS (see section 2.6.2).

2.7 Protein purification

2.7.1 Purification of AcsF$_{Ch}$

5 to 10 g frozen cells were resuspended in buffer A (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM TCEP) with a small amount of avidin and lysozyme. The cell suspension was sonicated (Bandelin Sonoplus 2200) three times for 5 min each (5 × 5 cycle, 50 % amplitude) in a rosette cooling cell on ice. It was then centrifuged in a polycarbonate bottle with cap assembly at 142 400 *g* for 1 hour at 12 °C. The supernatant was loaded on a streptactin column (10 - 30 mL) equilibrated in buffer A. After washing the column with 3 column volumes (CVs) of buffer A, AcsF$_{Ch}$ was eluted with a linear gradient of 0 - 2.5 mM desthiobiotin in the same buffer. The streptag of AcsF$_{Ch}$ was cleaved by incubation with a streptagged TEV protease for approximately 16 hours. The protein solution was applied to a PD10 desalting column and was then reloaded on the streptactin column equilibrated with buffer A. The flow-through was collected, concentrated and loaded on a Superdex 200 Hiload 16/60 column equilibrated in buffer B (50 mM Tris-HCl pH 8.0, 100 mM NaCl). Fractions corresponding to monomeric, dimeric and tetrameric AcsF$_{Ch}$ were pooled separately, concentrated, frozen in liquid N_2 (in glass vials equipped with a PTFE/silicone septum) and stored at -80 °C. For the variants K10A-AcsF$_{Ch}$ and C108A-AcsF$_{Ch}$, the purification was carried out using the same procedure.

2.7.2 Purification of AcsF$_{Af}$

AcsF$_{Af}$ was purified following the same protocol as described for AcsF$_{Ch}$ (see section 2.7.1), except that the streptag was not cleaved. After the streptactin column, the protein was directly loaded on the Superdex 200 size exclusion column. Fractions corresponding to monomeric and dimeric AcsF$_{Af}$ were pooled separately, concentrated and frozen in liquid N_2 and stored at -80 °C.

2.7.3 Purification of CooC3$_{Ch}$

CooC3$_{Ch}$ was purified following the same protocol as described for AcsF$_{Ch}$ (see section 2.7.1).

2.7.4 Purification of ACS

Cells containing the pET28a-ACSstrep or pET28a-ACStwinstrep plasmid were lysed, centrifuged and captured on the streptactin column as described for AcsF$_{Ch}$ in subsection 2.7.1. The protein was then eluted with desthiobiotin, concentrated and directly loaded on the Superdex 200 size exclusion column. The fractions corresponding to monomeric ACS were collected, concentrated, frozen in liquid N_2 and stored at -80 °C.

2.7.5 Sodium dodecyl sulfate polyacrylamide gel electrophoresis

To follow the course of a purification procedure or to analyze complex composition, the discontiniuous sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) as described by Laemmli was used [86]. The resolving gel was prepared by mixing either 12 % or 15 % polyacrylamide, 375 mM Tris-HCl pH 8.8, 0.1 % (w/v) sodium dodecyl sulfate (SDS), 0.08 % (v/v) N,N,N',N'-tetramethylethylenediamine (TEMED), and 0.05 % (w/v) ammonium persulfate (APS).The stacking gel contained 6 % acrylamide, 125 mM Tris-HCl pH 6.8, 0.1 % (w/v) SDS, and 0.08 % (v/v) TEMED.

The samples were mixed with an appropriate amount of 4x loading dye (200 mM Tris-HCl pH 6.8, 8 % (w/v) SDS, 0.4 % (w/v) bromophenol blue, 40 % (v/v) glycerol, 400 mM dithiothreitol), boiled for a few seconds at 100 °C and then loaded on the gel. The gel was run with a maximum voltage of 140 V. The Unstained Protein Molecular Weight Marker (Thermo Scientific), which ranges from 14.4 kDa to 116 kDa, was used as a molecular weight reference. The gel was stained overnight with Coomassie staining solution (0.025 % Coomassie Brilliant Blue G-250 (w/v), 10 % (v/v) acetic acid) and then destained with 10 % (v/v) acetic acid.

2.8 UV-Vis spectroscopy

UV-Vis spectra were obtained with an Agilent 8453 photodiode array spectrophotometer with a peltier temperature controller at 25 °C. Black walled quartz cuvettes with 1 cm pathlength were used to record spectra.

2.9 Acetyl-CoA formation activity of ACS

2.9.1 Synthesis of methylcobinamide

Methylcobinamide was synthesized from methylcobalamin as described previously by Zou et al. [87], except that, in this case, it was purified with a Sep Pak C18 reverse-phase column (Waters Sep Vak 6cc) and methylcobinamide was eluted with BuOH. Methylcobinamide was dried in an exicator over several days. It was then dissolved in Milli-Q H_2O and stored at -80 °C. All steps were carried out in darkness.

2.9.2 Reconstitution of apoACS

All incubation and reaction mixtures were set up inside an anoxic glove box (Model B; COY Laboratory Products). ApoACS was reconstituted by incubating 20 μM apoACS with 40 μM $NiCl_2$, 40 μM $AcsF_{Ch}$, 10 mM Mg-ATP in 0.1 M HEPES pH 7.2 at 60 °C. Four control experiments were set up; in each case, one component of the incubation mixture was missing. Aliquots were taken at different time points and the rate of acetyl-CoA formation was analyzed by UV-Vis spectroscopy.

2.9.3 Measurements of acetyl-CoA formation activity

Acetyl-CoA formation activity of ACS from CO, methylcobinamide and CoA was determined by following the conversion of methylcob(III)inamide (MeCbi) to cob(I)inamide (Cbi). The change of absorption was measured at 387 nm and 462 nm. The reaction rates were calculated using the differences of the extinction coefficients, $\Delta\varepsilon_{387nm\ (MeCbi\ -\ Cbi)} = -17410\ M^{-1}\ cm^{-1}$ and $\Delta\varepsilon_{465nm\ (MeCbi\ -\ Cbi)} = 7840\ M^{-1}\ cm^{-1}$, which were calculated

from spectra of Robertson et al. [88]. The reaction was set up inside the glove box as follows: 0.1 M HEPES pH 7.2, 50 µM MeCbi and 300 µM Ti(III)citrate were placed in a screw cap cuvette; the ACS incubation mixture (see section 2.9.2) was then added to yield a final concentration of 2 µM ACS. After closing the cap, CO-saturated buffer (10 % of the final volume) was injected with a syringe and the reaction was started by adding 200 µM CoA. All measurements were performed in darkness.

2.10 Measurements of NTPase activity

2.10.1 Coupled enzyme assay

To measure the NTPase activity of $AcsF_{Ch}$, a coupled assay using pyruvate kinase (PK) and lactate dehydrogenase (LDH) was performed [89]. In this assay the conversion of one molecule of NADH to NAD^+ by lactate dehydrogenase equals the consumption of one molecule of ATP by $AcsF_{Ch}$. The rate of ATP hydrolysis at 25 °C was determined by measuring the decrease of absorption at 340 nm. Reaction rates were calculated using the extinction coefficient of NADH ε_{340nm} = 6220 M^{-1} cm^{-1}. The assay solution was prepared in 50 mM HEPES pH 7.2 containing 150 µM KCl, 10 mM $MgCl_2$, 2.3 mM phosphoenolpyruvate, 140 µM NADH, 9 units LDH, 6 units PK and 1 mM ATP. The assay was started by adding 5 µM $AcsF_{Ch}$.

2.10.2 Malachite green assay

The malachite green assay was performed at 45 °C and 60 °C to determine the rate of ATP hydrolysis at elevated temperatures. In this assay the amount of phosphate released by the hydrolysis of ATP is detected by following the increase of absorption at 630 nm [90]. This increase is due to the formation of a malachite green phosphomolybdate complex. The rate calculation was based on a standard curve, which was prepared with KH_2PO_4 in the range from 1 µM to 9 µM. The assay solution contained 5 µM $AcsF_{Ch}$, 1 mM Mg-ATP in 0.1 mM HEPES pH 7.2. Aliquots were taken at different time points and mixed with the malachite green ammonium molybdate solution. A solution of 34 % (w/v) sodium

citrate (10 % of the final volume) was immediately added. This mixture was incubated for 15 min and the absorption was measured at 630 nm. Control experiments in the absence of $AcsF_{Ch}$ were performed to assess the rate of autohydrolysis of ATP at 60 °C.

2.11 Oligomerization and complex formation behaviour

2.11.1 Calibration of the size exclusion column

A calibration of the Superdex 200 HiLoad 16/60 column equilibrated in 50 mM Tris-HCl pH 8.0, 100 mM NaCl was performed with the following molecular weight standards: aprotinin (6.5 kDa), RNase A (13.7 kDa), DNase I (31 kDa), ovalbumin (44 kDa), conalbumin (75 kDa), lactate dehydrogenase (140 kDa), aldolase (161 kDa) and catalase (240 kDa). The decadic logarithm of the molecular weight was plotted against the elution volume and the calibration curve was determined by linear regression. For the calibration and the following experiments, the flow rate was 1 mL min^{-1}. The elution profiles were recorded by following the absorption at 280 nm.

2.11.2 Analysis of the dimerization behaviour of $AcsF_{Ch}$ in the presence of nucleotide

The dimerization behaviour of $AcsF_{Ch}$ in the presence of nucleotide was analyzed by size exclusion chromatography. Monomeric $AcsF_{Ch}$ was incubated in 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM TCEP, 0.25 mM $MgCl_2$ and 0.25 mM nucleotide for 10 min. The nucleotide was either ATP or ADP. The incubation mixture was then loaded on a Superdex 200 HiLoad 16/60 column equilibrated with the incubation buffer containing the corresponding nucleotide.

2.11.3 Analysis of the complex formation behaviour of ACS with $AcsF_{Ch}$, K10A-$AcsF_{Ch}$, C108A-$AcsF_{Ch}$, $CooC1_{Ch}$ and $CooC3_{Ch}$

ApoACS and $AcsF_{Ch}$ were incubated for 10 min inside the glove box and loaded on a Superdex 200 HiLoad 16/60 column equilibrated in 50 mM Tris-HCl pH 8.0, 100 mM NaCl. As a control apoACS and $AcsF_{Ch}$ were also loaded individually on the column.

From these three chromatography runs, the fractions were analyzed by SDS-PAGE. The stoichiometry of the complex was analyzed with the software GelAnalyzer 2010a [91]. The complex formation behaviour of apoACS with K10A-AcsF$_{Ch}$, C108A-AcsF$_{Ch}$, CooC1$_{Ch}$ and CooC3$_{Ch}$ was also analyzed by size exclusion chromatography as described for wildtype AcsF$_{Ch}$.

2.12 Metal binding analysis

2.12.1 UV-Vis spectroscopic nickel binding assay

20 µM AcsF$_{Ch}$ in 50 mM Tris-HCl pH 8.0, 100 mM NaCl were titrated with $NiCl_2$ to yield concentrations of 1 µM, 2 µM, 4 µM, 6 µM, 8 µM, 10 µM, 12 µM, 14 µM, 16 µM, 18 µM, 20 µM and 40 µM Ni^{2+}. After each addition one spectrum was recorded. Titrations and UV-Vis spectroscopy were performed inside the glove box.

2.12.2 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments were performed with a MicroCal VP-ITC system (GE) at 25 °C in 50 mM Tris-HCl, 100 mM NaCl. Three different experiments were performed: a) 15 µM AcsF$_{Ch}$ were titrated with 300 µM $NiCl_2$, b) 24.7 µM apoACS were titrated with 750 µM $NiCl_2$ and c) 24.4 µM ACS-AcsF$_{Ch}$ complex were titrated with 750 µM $NiCl_2$. The concentrations of apoACS and ACS-AcsF$_{Ch}$ complex were normalized to the amount of iron, which was determined after ITC measurements by inductively coupled plasma atomic emission spectroscopy. The enzyme was present in the reaction cell (1.4 mL) and $NiCl_2$ was in the syringe. The initial injection was 3 µL (lasting 6 s) and the 47 subsequent injections were 6 µL (lasting 12 s). There was an interval of 300 s between injections. The solution in the reaction cell was stirred at 307 rpm. The integration of the peak areas of the time dependent raw data was done with NITPIC [92]. This data was imported to the Origin 7 software provided by the manufacturer and fitted to a one-site model. The Wiseman “c” value is defined as the product of the protein concentration in the cell and the binding constant [93].

2.12.3 Removal of excess metal after ITC experiments

When the ITC experiment was finished, samples b) and c) were used for further experiments, i.e. activity measurements and metal analysis. The protein samples were concentrated to approximately 300 µL and the excess metal was removed using a PD MiniTrap-G25 column. Only coloured fractions were collected. To ensure that all excess metal was removed a buffer sample was prepared for metal analysis: 130 µM $NiCl_2$ were added to the buffer and the sample was treated exactly the same as the protein samples.

2.12.4 Metal analysis by inductively coupled plasma optical emission spectroscopy

Metal analysis was performed by Jasmin Kurtzke from the group of Prof. Silke Leimkühler at the University of Potsdam. She used the following protocol: 500 µL protein samples were wet-ashed overnight in a 1:1 mixture with 65 % nitric acid (Suprapur, Merck, Darmstadt, Germany) at 100 °C. The samples were diluted with 4 mL of H_2O prior to their injection onto a PerkinElmer Optima 2100 DV inductively coupled plasma optical emission spectroscopy (ICP-OES) system (Fremont, CA, USA). The multielement standard solution XVI (Merck) was used as a reference.

2.13 Sequence similarity analysis and alignment visualization

Sequences of the CooC and AcsF groups were aligned with MAFFT [94, 95]. The Kimura distance matrix was calculated by Prof. Holger Dobbek using protdist/PHYLIP [96]. The sequence similarity network was visualized with Cytoscape 3.3.0 [97]. The MAFFT alignment was also imported into Alvis [98] and visualized as sequence bundles.

2.14 Protein crystallization and structure determination

Crystallization experiments were performed under anoxic conditions at 18 °C. For the initial screening of crystallization conditions, the sitting drop vapor diffusion technique was used. Drops were set up with the OryxNano crystallization robot (Douglas Instruments) in 96-well plates (Greiner). The drop size was 1 - 1.5 µL and the reservoir volume was 90 µL. The ratio of protein solution and reservoir solution was 1:1. The initial screens were performed with commercially available crystallization screens. Suitable conditions were optimized with the hanging drop vapor diffusion technique and set up in 24-well plates (cellstar, greiner bio-one) with a drop size of 4 µL and a reservoir volume of 600 µL. Within these 24 conditions, the buffer, the pH, the precipitant concentration and the protein concentration were systematically varied. Crystals were fished with cryoloops, quickly soaked with a cryoprotective solution (85 % reservoir solution and 15 % 1,4-butanediol) and then flash frozen in liquid nitrogen.

2.14.1 Data collection and processing

Diffraction images of the crystals were collected at 100 K on beamline BL 14.1 or BL 14.2 at the synchrotron radiation source BESSY II (Helmholtz Zentrum Berlin) [99]. The measurement wavelength was 0.9184 Å if not otherwise noted. The optimal data collection strategy was determined with the program iMOSFLM [100]. The diffraction data sets were indexed, integrated and scaled with XDS and XDSAPP [101, 102].

2.14.2 Crystallization and data collection of CooC3$_{Ch}$

Dr. Jae-Hun Jeoung purified and crystallized CooC3$_{Ch}$. The protein was not expressed and purified as described in section 2.7.3, but it was expressed with a pET21b vector and purified by Ni-immobilized metal affinity chromatography. CooC3$_{Ch}$ (10 mg mL^{-1}) was preincubated with 1 mM $NiCl_2$ and crystallized in a condition containing 0.2 M $(NH_4)_2SO_4$, 0.1 M Tris-HCl pH 8.5 and 22 % PEG 3350. Butandiole was used as a cryo-

protectant. The dataset was collected at 1.47 Å. The dataset was kindly provided by Dr. Jae-Hun Jeoung to be used as part of this project.

2.14.3 Structure determination

The structures were determined by molecular replacement, which was carried out with PHENIX Phaser [103]. The search model "short-CooC1" was prepared with PHENIX Sculptor [104]. The search model "short-AcsF" was prepared manually. Initial models were built with PHENIX AutoBuild [105]. The model was completed by iterative cycles of manual model building with COOT [106] and refinement with PHENIX Refine [107].

2.14.4 Structure representation

Molecular structure representations were prepared with PyMOL [108]. Topology diagrams were prepared with Pro-origami [109] and adapted with Adobe Illustrator.

3 | Results and discussion

3.1 Protein expression and purification

3.1.1 Expression and purification of AcsF$_{Ch}$

The expression of the gene encoding for AcsF$_{Ch}$ in *E. coli* was improved by changing the vector system from a pASK-IBA17plus vector to a pET28aTEVstrep vector (Fig. A.4 in the Appendices). The cultivation of *E. coli* BL21 with this plasmid yielded approximately 4 g of wet cell paste per liter of LB medium. AcsF$_{Ch}$ and the variants K10A-AcsF$_{Ch}$ and C108A-AcsF$_{Ch}$ were purified in a three-step procedure: 1) affinity chromatography, 2) a second passage over the affinity column after cleavage of the tag by TEV protease and 3) size exclusion chromatography to separate different oligomeric states of AcsF$_{Ch}$. Figure 3.1 A shows a typical size exclusion chromatogram, which has two peaks at 83.2 mL and 90.5 mL. A standard curve (Fig. A.5 in the Appendices) was used to determine the corresponding molecular weights, 58 kDa and 29 kDa, respectively. These values agree with the predicted molecular weights of dimeric (55.4 kDa) and monomeric AcsF$_{Ch}$ (27.7 kDa). Occasionally, a third peak in the chromatogram was observed at 76 mL (Fig. 3.1, dotted line). The calculated molecular weight of this species is 114.6 kDa, corresponding to tetrameric AcsF$_{Ch}$ (115 kDa). This fraction, however, was not active in the activation assay (see section 3.7), and was not investigated further. The ratio of the oligomeric states seemed to depend on the cultivation batch, rather than on individual purifications. The variant K10A-AcsF$_{Ch}$, in which the signature lysine of the deviant Walker A motif is replaced by alanine, was present in monomeric and dimeric states. By contrast, the C108A-AcsF$_{Ch}$ variant, in which the first of the cysteine residues

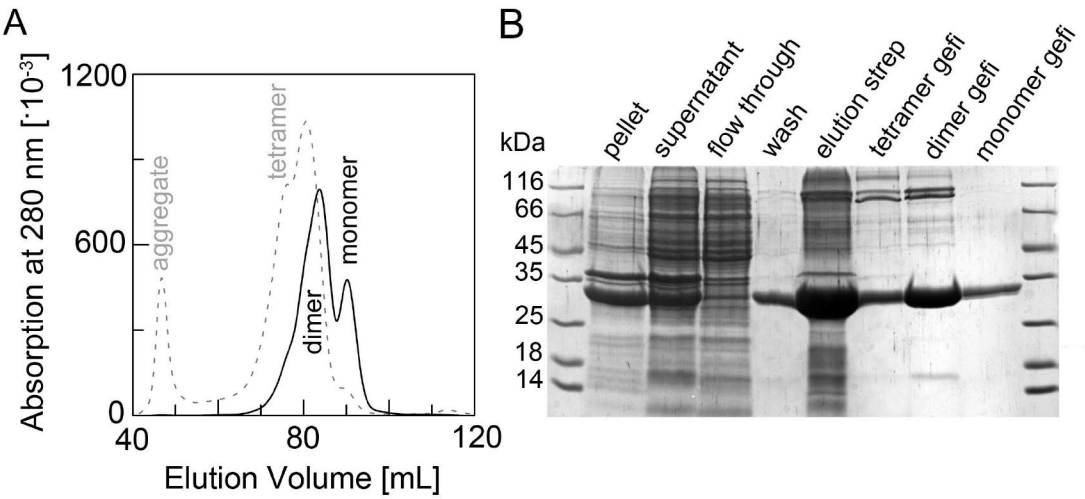


Figure 3.1: Purification of AcsF$_{Ch}$. A) The black solid line depicts a typical size exclusion chromatogram with peaks for monomeric and dimeric AcsF$_{Ch}$. The grey dotted line depicts an atypical chromatogram with an additional peak for tetrameric AcsF$_{Ch}$. B) SDS-PAGE gel of the purification of AcsF$_{Ch}$. The gel was loaded with samples from various stages of the purification: with samples of the pellet and supernatant; with samples from the flow through, the wash fraction and the eluate of the streptactin column; and finally with samples from the tetrameric, dimeric and monomeric fractions from the size exclusion column. The first and last lane contained a marker with the indicated molecular masses. Gefi stands for gelfiltration.

of the conserved CXC motif was replaced by alanine, was present only in the monomeric state. The SDS gel in Figure 3.1 B shows the course of the purification. The final purity of AcsF$_{Ch}$ was higher than 96 %, as determined by SDS densitometric analysis.

3.1.2 Expression and purification of AcsF$_{Af}$

AcsF$_{Af}$ was purified in two stages: an affinity chromatography followed by a size exclusion chromatography. In contrast to the purification procedure of AcsF$_{Ch}$, the streptag of AcsF$_{Af}$ was not cleaved as AcsF$_{Af}$ tended to precipitate during incubation with TEV protease. AcsF$_{Af}$ was present in both the monomeric and dimeric state, eluting at 89 mL and 80 mL, respectively. Compared to AcsF$_{Ch}$, the elution volume of the dimeric state is shifted to slighlty higher values even though the molecular weight differs

only by 3.9 kDa. This could indicate a more elongated shape of the dimer compared to $AcsF_{Ch}$. The SDS gel in Figure 3.2 shows the course of the purification. The purity of $AcsF_{Af}$, which was determined by densitometric analysis was more than 98 %.

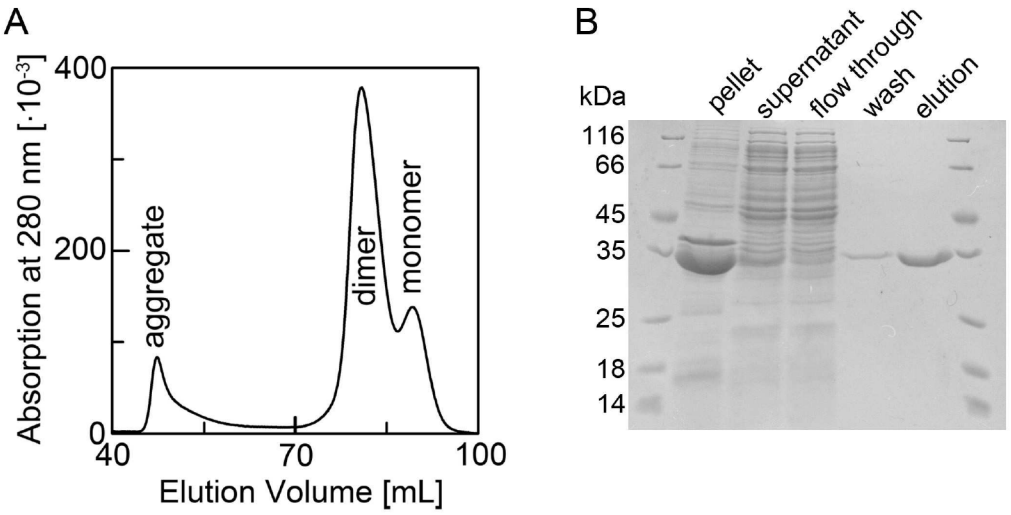


Figure 3.2: Purification of $AcsF_{Af}$. A) Typical size exclusion chromatogram of the purification of $AcsF_{Af}$. Most of $AcsF_{Af}$ is present in the dimeric state. B) SDS-PAGE gel of the purification of $AcsF_{Af}$. The gel was loaded with samples of the pellet, the supernatant and three samples from the streptag affinity chromatography (the flow through, the wash fraction and the eluate). The first and last lane contained a marker with the indicated molecular masses. A large fraction of $AcsF_{Af}$ seems to be unsoluble, as the band for $AcsF_{Af}$ is much more pronounced in the lane containing the pellet than in the lane containing the supernatant.

3.1.3 Expression and purification of ACS

Anaerobic cultivation of *E. coli* harbouring the pET28a-ACS(twin)strep plasmid in 5 L bottles yielded a cell mass of approximately 1.5 g per liter of mTB medium. ACS was purified in two steps: an affinity chromatography followed by a size exclusion chromatography. One issue encountered during the purification of ACS with an N-terminal streptag was that it did not bind very well to the column. This was overcome by switching the tag from a single streptag to a twin-streptag, which improved binding considerably. A potential explanation for this is that the single streptag was shielded by the protein

and did not have access to the streptactin resin. The twin-streptag is longer and more flexible, and has therefore better access to the streptactin binding pocket. Changing the tag also altered the size exclusion chromatogram (Fig. 3.3 A). It reduced the presence of aggregated ACS, dimeric ACS and smaller fragments of ACS and shifted the equilibrium to monomeric ACS, which eluted at 76 mL. According to the standard curve this corresponds to a molecular weight of 115 kDa, which is considerably larger than the calculated molecular weight of 86.6 kDa. This suggests that ACS does not have a perfect globular shape in solution. The SDS gel, which shows the course of the purification is shown in Figure 3.3 B. The final purity of ACS was approximately 92 % as determined by densitometric analysis.

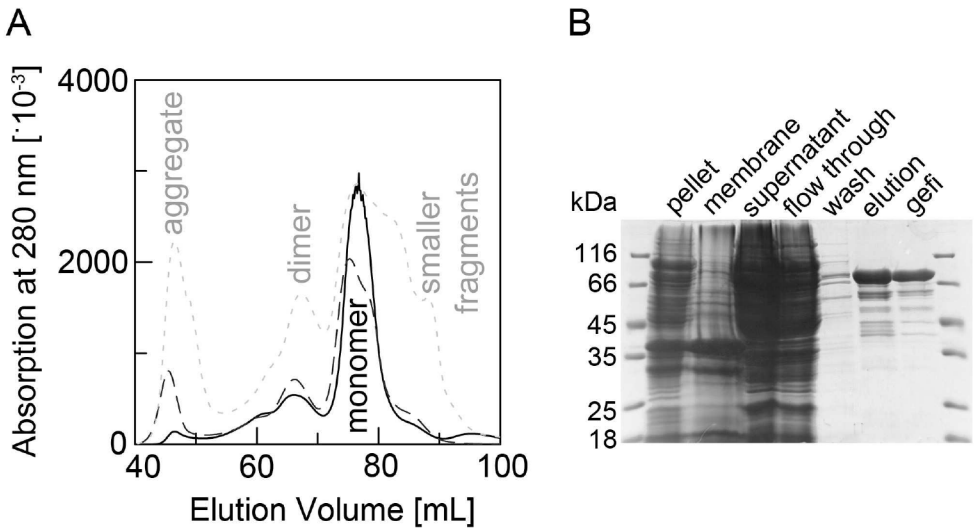


Figure 3.3: Purification of ACS. A) Size exclusion chromatograms of ACS. A typical size exclusion chromatogram of the purification of ACS with a twin-streptag is shown as black solid line. The dashed and dotted lines in grey depict chromatograms of ACS that contained a streptag. B) SDS-PAGE gel of the purification of ACS. The gel was loaded with samples of various stages of the purification: with samples from the pellet, the membrane layer and the supernatant; samples from the flow through, the wash fraction and the eluate of the streptag column; and finally with a sample of the monomeric fraction of the size exclusion chromatography. The first and last lane contained a marker with the indicated molecular masses.

3.2 UV-Vis spectral features of AcsF$_{Ch}$ and ACS

The UV-Vis spectrum of AcsF$_{Ch}$ is shown in Figure 3.4. The maximum of absorption is at 274 nm rather than at 280 nm, because AcsF$_{Ch}$ does not contain tryptophan residues. The UV-Vis spectrum of as-isolated ACS (Fig. 3.4) contains a long shoulder around 420 nm, which is consistent with the presence of a $[Fe_4S_4]^{2+}$ cluster. Overall, the spectrum of as-isolated ACS from heterologous expression in *E. coli* has the same features as the spectrum of ACS purified from *C. hydrogenoformans*, which was obtained by Svetchlitchny and colleagues [7].

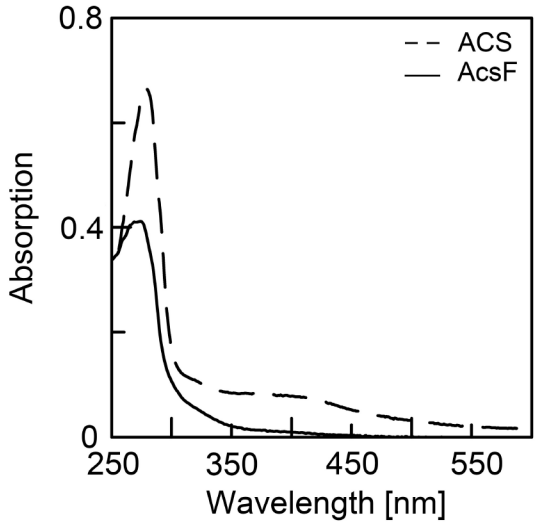


Figure 3.4: UV-Vis spectrum of AcsF$_{Ch}$ and ACS. The spectrum of AcsF$_{Ch}$ is shown as solid line and the spectrum of ACS is shown as dashed line.

3.3 Nickel binding behaviour of AcsF$_{Ch}$

3.3.1 Spectroscopic nickel binding assay

In light of the proposed involvement of AcsF$_{Ch}$ in the maturation of ACS, along with the presence of a potential metal binding motif (CXC), we decided to investigate whether or not AcsF$_{Ch}$ binds nickel. For other maturation proteins of nickel-containing enzymes,

e.g. CooC1$_{Ch}$ or HypB, it has been shown, that upon nickel addition, the absorption increases in the region between 310 nm and 350 nm [64, 110]. This region is characteristic for a ligand to metal charge transfer (LMCT) between Cys-S$^-$ to Ni(II). If AcsF$_{Ch}$ did bind nickel, one would expect similar spectral changes. Upon addition of increasing amounts of $NiCl_2$ to AcsF$_{Ch}$, however, there was no distinct spectral change (Fig. 3.5 A). Instead, there was a general rise of absorption, which was due to precipitation of AcsF$_{Ch}$. Moreover, the same experiment was carried out in the presence of nucleotides, but no new spectral features developed upon addition of $NiCl_2$ (data not shown).

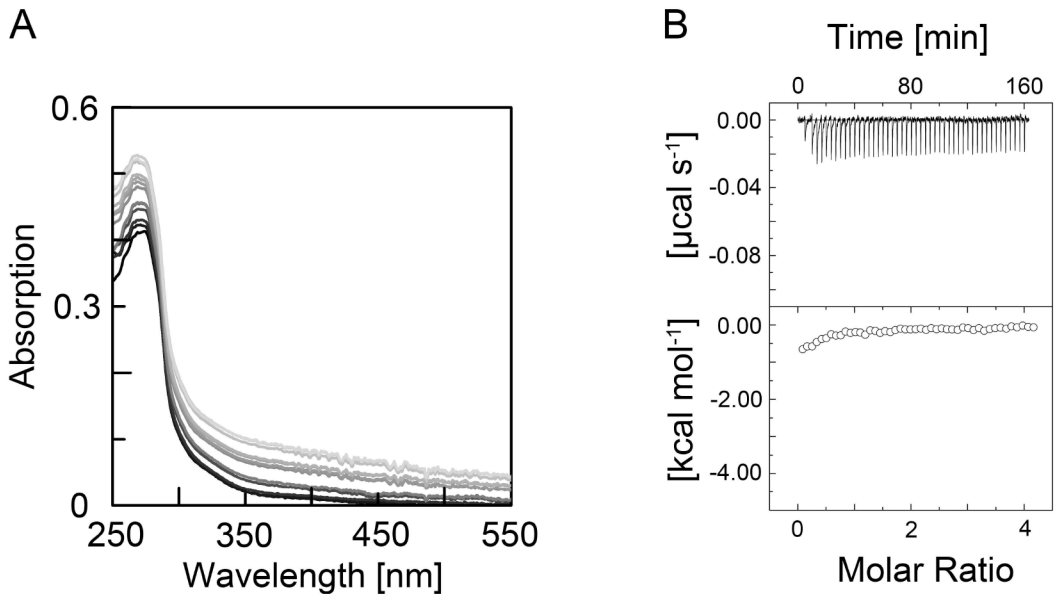


Figure 3.5: A) Titration of AcsF$_{Ch}$ with $NiCl_2$. The spectrum of 20 µM AcsF$_{Ch}$ is shown as a black line. The same AcsF$_{Ch}$ sample was then titrated with $NiCl_2$ to yield molecular concentrations of 1 µM, 2 µM, 4 µM, 6 µM, 8 µM, 10 µM, 12 µM, 14 µM, 16 µM, 18 µM, 20 µM and 40 µM Ni^{2+}. For each concentration one spectrum was recorded. As the amount of nickel increases, the spectra of AcsF$_{Ch}$ are shown in progressively lighter shades of grey. No new spectral features developed upon titration of nickel to AcsF$_{Ch}$. B) Isothermal titration calorimetry of AcsF$_{Ch}$ with $NiCl_2$. The top panel shows the ITC raw data. The bottom panel shows the normalized heat change of injections, which was calculated by integration of the peak areas of the raw data. There was no significant enthalpy change upon titration of AcsF$_{Ch}$ with $NiCl_2$

3.3.2 Isothermal titration colorimetry of $AcsF_{Ch}$ with nickel

Furthermore, an ITC experiment was carried out to test whether $AcsF_{Ch}$ binds nickel. This experiment confirmed the results from the UV-Vis spectroscopy. Upon titration of $AcsF_{Ch}$ with $NiCl_2$ only heats of dilution were observed and no binding event was detected (Fig. 3.5 B). Both experiments showed that $AcsF_{Ch}$ is not able to bind nickel in the as-isolated state.

3.4 Dimerization behaviour of $AcsF_{Ch}$ in the presence of nucleotides

According to the proposed mechanism (see section 1.4) members of the MinD family must undergo dimerization to hydrolyze ATP. To test whether this is also true for $AcsF_{Ch}$, its oligomeric state was analyzed by size exclusion chromatography in the presence of either Mg-ADP or Mg-ATP (Fig. 3.6 A). In the presence of Mg-ADP, the elution volume did not change compared to monomeric $AcsF_{Ch}$ showing that Mg-ADP has no effect on the oligomeric state of $AcsF_{Ch}$. This is in contrast to the behaviour of $CooC1_{Ch}$, which dimerized in the presence of Mg-ADP [64]. The presence of Mg-ATP, however, lead to the dimerization of $AcsF_{Ch}$ as predicted from the proposed mechansim.

To evaluate the role of the signature lysine in $AcsF_{Ch}$, the same experiment was performed with the K10A-$AcsF_{Ch}$ variant. K10A-$AcsF_{Ch}$ did not dimerize, neither in the presence of Mg-ADP nor in the presence of Mg-ATP (Fig. 3.6 B). These results highlight that the signature lysine residue plays a critical role in the function of $AcsF_{Ch}$.

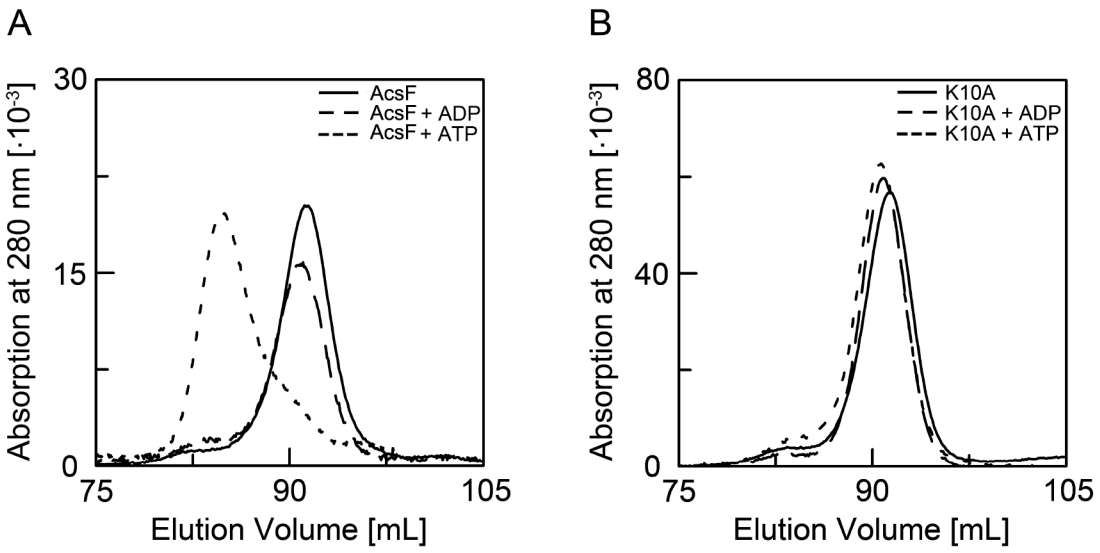


Figure 3.6: A) Nucleotide dependent dimerization of $AcsF_{Ch}$. The elution profile of monomeric $AcsF_{Ch}$ without nucleotide is shown in solid line and the elution profiles of $AcsF_{Ch}$ in the presence of Mg-ADP and Mg-ATP are shown in dashed and dotted lines, respectively. The presence of Mg-ADP in the buffer did not change the oligomeric state of $AcsF_{Ch}$. In the presence of Mg-ATP, however, the peak shifted by 6 mL from monomeric to dimeric $AcsF_{Ch}$. B) Influence of nucleotide on the oligomeric state of the variant K10A-$AcsF_{Ch}$. The elution profile of monomeric K10A-$AcsF_{Ch}$ is shown in solid line, the elution profile of K10A-$AcsF_{Ch}$ in the presence of Mg-ADP is shown in dashed line and the elution profile of K10A-$AcsF_{Ch}$ in the presence of Mg-ATP is shown in dotted line. The variant K10A-$AcsF_{Ch}$ did not form a dimer in the presence of nucleotide.

3.5 ATPase activity of $AcsF_{Ch}$

The measured activities varied slightly depending on the purification batch. The average activities measured were 6.2 nmol min^{-1} mg^{-1} at 25 °C, 23.5 nmol min^{-1} mg^{-1} at 45 °C and 56.5 nmol min^{-1} mg^{-1} at 60 °C (Table 3.1). $AcsF_{Ch}$ did not hydrolyze ATP in the absence of magnesium. This is in line with the proposal that magnesium stabilizes the negative charges during ATP hydrolysis. Crystal structures of MinD proteins showed that the magnesium ion is typically coordinated to residues of the switch I motif and the

β-phosphate of the nucleotide [111, 112]. In addition, these measurements illustrate the temperature stability of AcsF$_{Ch}$. Even at 60 °C, AcsF$_{Ch}$ is stable and hydrolyzes ATP.

The variant K10A-AcsF$_{Ch}$, however, does not hydrolyze ATP (Table 3.1). This agrees with the current working model of MinD-type ATPases, which presumes that the signature lysine is an invariant residue crucial to the function of the protein.

The variant K10A-AcsF$_{Ch}$ and wildtype AcsF$_{Ch}$ were purified following the same protocol. One can therefore conclude that the activities measured for wildtype AcsF$_{Ch}$ are unlikely to originate from impurities, since if they did, similar values would be expected for the variant K10A-AcsF$_{Ch}$. Taken together, these measurements show that AcsF$_{Ch}$ is an ATPase of low specific activity.

Table 3.1: ATPase activity of AcsF$_{Ch}$ and the variant K10A-AcsF$_{Ch}$. Different purification batches are numbered #1 - #7. Each measurement was carried out three times, except for the NADH stability test, which was measured once. n.d. stands for not detected.

	specific activity [nmol min^{-1} mg^{-1}]				
	coupled assay		malachite green assay		
	25 °C	45 °C	25 °C	45 °C	60 °C
#1 AcsF$_{Ch}$	5.7 ± 0.2	29.4 ± 5.3			
#2 AcsF$_{Ch}$			4.6 ± 0.8	26.5 ± 1.2	69.1 ± 2.4
#3 AcsF$_{Ch}$			3.2 ± 0.2	14.3 ± 0.2	43.9 ± 2.5
#4 AcsF$_{Ch}$	8.1 ± 0.4				
#5 AcsF$_{Ch}$		19.8 ± 2.4			
#6 AcsF$_{Ch}$	10.2 ± 3.0				
#7 AcsF$_{Ch}$	5.5 ± 0.5	27.4 ± 4.0			
#1 K10A-AcsF$_{Ch}$	n.d.	2.5 ± 1.3			
#2 K10A-AcsF$_{Ch}$	2.4 ± 0.1	2.8 ± 0.5			
ATP autohydrolysis					n.d.
NADH stability	0.88	2.11			
ATP autohydrolysis (+ NADH stability)	1.5 ± 0.5	2.2 ± 0.2			

3.6 Complex formation behaviour of ACS

3.6.1 Complex formation of apoACS with AcsF$_{Ch}$

To establish whether or not AcsF$_{Ch}$ interacts with ACS, a size exclusion chromatography was performed with a 1:2 mixture of apoACS and AcsF$_{Ch}$. The chromatogram showed a new peak at 70.2 mL, clearly distinct from the peaks of apoACS and AcsF$_{Ch}$, which appear at 76.0 and 83.2 mL, respectively (Fig. 3.7 A). The molecular weight of the new species, which was calculated according to the standard curve, is 199 kDa. Since apoACS and AcsF$_{Ch}$ were incubated in a 1:2 stoichiometry and no other peaks were apparent in the chromatogram, one can conclude that apoACS and AcsF$_{Ch}$ formed a complex with a stoichiometry of 1:2. The discrepancy between the calculated (199 kDa) and predicted molecular weight (141 kDa) might be due to a non-globular shape of the complex, which would lead to a longer retention time. SDS-PAGE of the peak fraction of the new species showed two bands at the predicted sizes for apoACS and AcsF$_{Ch}$, which clearly demonstrates that AcsF$_{Ch}$ can form a complex with apoACS (Fig. 3.7 B). A detailed analysis of the SDS-PAGE gel with the software GelAnalyzer 2010a also confirmed that the stoichiometry of the complex is one molecule of apoACS and two molecules of AcsF$_{Ch}$. Complex formation occurred regardless of whether apoACS was incubated with dimeric or monomeric AcsF. Interestingly, the presence of Mg-ADP, Mg-ATP or $NiCl_2$ had no effect on the complex formation behaviour.

3.6.2 Complex formation of ACS with apoAcsF$_{Ch}$ variants and CooC proteins

To assess the specificity of the interaction between apoACS and AcsF$_{Ch}$, it was also investigated whether or not apoACS forms a complex with CooC1$_{Ch}$ or CooC3$_{Ch}$. These proteins are involved in the maturation of CODH and share 28.6 % and 28.3 % sequence identity, respectively, with AcsF$_{Ch}$ (Fig. A.6 in the Appendices). Both proteins also contain the CXC motif, which is crucial to the interaction of AcsF$_{Ch}$ and apoACS (see paragraph below). Despite these similarities, neither CooC1$_{Ch}$ nor CooC3$_{Ch}$ forms a stable

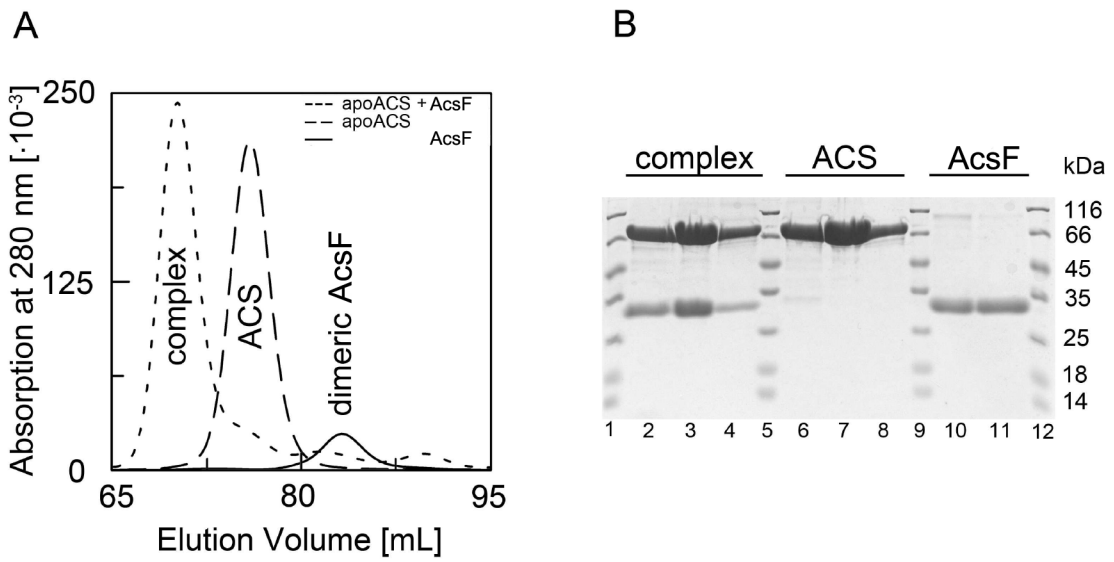


Figure 3.7: Complex formation of apoACS with $AcsF_{Ch}$. A) Size exclusion chromatograms of dimeric $AcsF_{Ch}$, apoACS and the apoACS-$AcsF_{Ch}$ complex. The apoACS-$AcsF_{Ch}$ complex was prepared by incubating 13 nmol apoACS with 13 nmol dimeric $AcsF_{Ch}$ for 10 min. The elution profile of the apoACS-$AcsF_{Ch}$ complex is shown in dotted line and the elution profiles of the individual proteins apoACS and $AcsF_{Ch}$ are shown in dashed and solid lines, respectively. B) SDS" PAGE analysis of size exclusion chromatography experiments. From each chromatography run, the main peak fractions were loaded on a gel (apoACS-$AcsF_{Ch}$ complex: 67 mL to 76 mL, apoACS: 72.5 mL to 80 mL, $AcsF_{Ch}$: 80.5 mL to 85.5 mL). The first, fifth, ninth and last lanes show a marker with the indicated molecular masses.

complex with apoACS (Fig. 3.8 A). This demonstrates that the interaction of $AcsF_{Ch}$ and apoACS is specific.

Furthermore, it was determined whether the variants K10A-$AcsF_{Ch}$ and C108A-$AcsF_{Ch}$ also form a complex with apoACS (Fig. 3.8 B). The size exclusion chromatogram shows that the replacement of the signature lysine residue with an alanine residue does not influence complex formation, whereas the substitution of one of the cysteine residues disrupts complex formation.

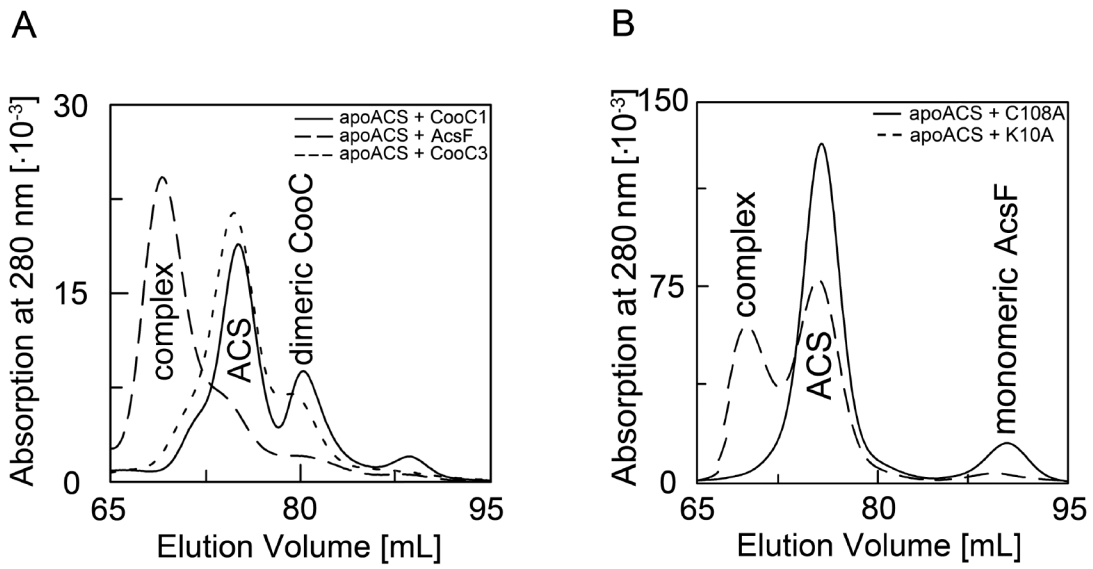


Figure 3.8: Complex formation of apoACS with CooC-type proteins and AcsF$_{Ch}$ variants. A) Size exclusion chromatograms of incubation mixtures of apoACS with CooC1$_{Ch}$, CooC3$_{Ch}$ and AcsF$_{Ch}$ shown in solid, dotted and dashed line, respectively. In contrast to AcsF$_{Ch}$, the CooC-type proteins CooC1$_{Ch}$ and CooC3$_{Ch}$ do not form a stable complex with apoACS. B) Size exclusion chromatograms of incubation mixtures of apoACS with the variants K10A-AcsF$_{Ch}$ and C108A-AcsF$_{Ch}$ shown in dashed and solid line, respectively. The incubation mixtures consisted of either monomeric K10A-AcsF$_{Ch}$ or monomeric C108A-AcsF$_{Ch}$ and an excess amount of apoACS. The variant C108A-AcsF$_{Ch}$ is not able to form a complex with ACS.

3.7 Reconstitution of ACS

3.7.1 Reconstitution of apoACS by incubation with AcsF$_{Ch}$, Mg-ATP and $NiCl_2$

When bacterial ACS or the β-subunit of the archaeal multisubunit ACDS complex are heterologously expressed in *E. coli*, the enzyme is not active. This is due to the incomplete assembly of the A-cluster of ACS in *E. coli*. Analysis of the metal content suggests that the [Fe_4S_4]-component of the cluster is present after purification, but the two nickel ions are lacking (see section 3.8). The [Fe_4S_4]-component is most likely assembled by the *E. coli* Isc machinery [113]. *E. coli*, however, does not seem to contain proteins which can insert nickel into ACS. Acetyl-CoA formation activity of ACS can be restored

by incubating ACS with an excess amount of nickel [43, 114, 115]. The lack of nickel is usually a complicating factor, but in our case it served as an optimal starting point from which to investigate the maturation of ACS.

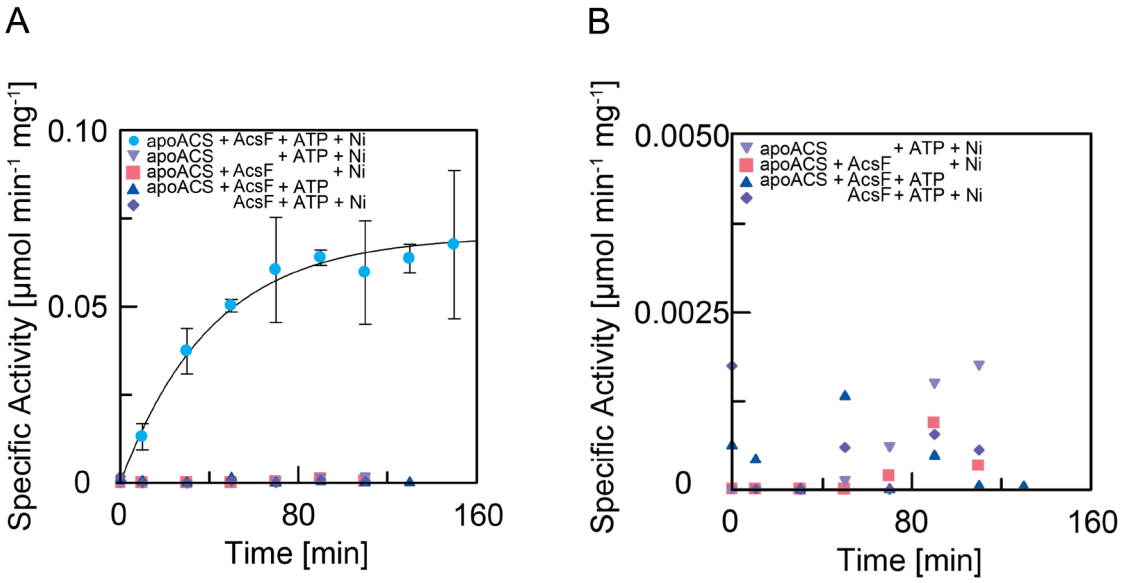


Figure 3.9: A) Time dependent activation of apoACS in the presence of $AcsF_{Ch}$, Mg-ATP and $NiCl_2$. 20 µM apoACS were incubated with 20 µM dimeric $AcsF_{Ch}$, 10 mM Mg-ATP and 40 µM $NiCl_2$ at 60 °C. Aliquots of incubation mixtures were taken at indicated time points and the specific activity was measured by following the conversion of Mecob(III)inamid to cob(I)inamid by UV-Vis spectroscopy. Each data point was measured in triplicate and error bars represent the standard deviation. Four control experiments were performed, in which either $AcsF_{Ch}$, Mg-ATP, $NiCl_2$, or apoACS was omitted from the incubation mixture. Data points for the control experiments were measured once. The activation of apoACS with stoichiometric amounts of nickel is only possible in the presence of both $AcsF_{Ch}$ and Mg-ATP simultaneously. B) Magnification of the data points of the control reactions.

We performed a reconstitution assay at 60 °C, in which only stoichiometric amounts of nickel were added to apoACS in the presence of $AcsF_{Ch}$ and Mg-ATP (Fig. 3.9). Within two hours, ACS gained activity, i.e. 0.7 µmol min^{-1} mg^{-1} at 25 °C. In the absence of $AcsF_{Ch}$, however, stoichiometric amounts of nickel were not sufficient to activate apoACS within this time frame. The activation of ACS by $AcsF_{Ch}$ is strictly ATP dependent, as there

was no gain in activity when Mg-ATP was omitted from the incubation mixture. These results clearly show that AcsF$_{Ch}$ is involved in the maturation of the A-cluster of ACS and that this maturation is Mg-ATP-dependent.

3.7.2 Reconstitution of ACS by apoAcsF$_{Ch}$ variants and CooC proteins

It was also investigated how the amino acid exchanges K10A and C108A influence the capability of AcsF$_{Ch}$ to activate apoACS. Neither the variant K10A-AcsF$_{Ch}$, nor the variant C108A-AcsF$_{Ch}$ was able to activate apoACS (Fig. 3.10). This indicates that the ATPase activity of AcsF$_{Ch}$, as well as the presence of the CXC motif, are essential for the activation process.

Furthermore it was investigated whether CooC1$_{Ch}$ or CooC3$_{Ch}$ could act as a maturation factor for ACS. Neither CooC1$_{Ch}$ nor CooC3$_{Ch}$ can replace AcsF$_{Ch}$ in the activation assay (Fig. 3.10). This shows that AcsF$_{Ch}$ is a specific maturation factor of ACS.

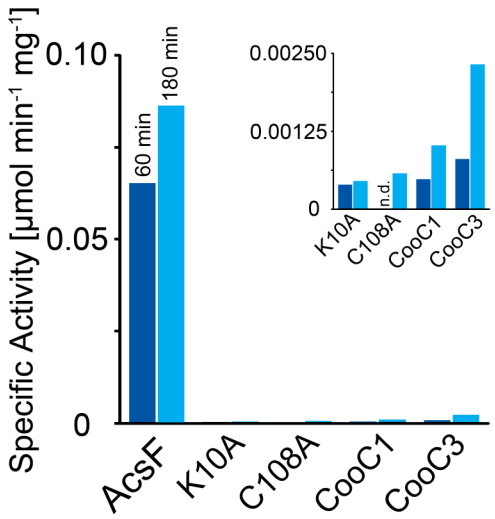


Figure 3.10: Activity of ACS that was reconstituted with AcsF$_{Ch}$, K10A-AcsF$_{Ch}$, C108A-AcsF$_{Ch}$, CooC1$_{Ch}$ or CooC3$_{Ch}$ for 60 min (dark blue) and 180 min (light blue).The inset shows a magnification of the data points for K10A-AcsF$_{Ch}$, C108A-AcsF$_{Ch}$, CooC1$_{Ch}$ and CooC3$_{Ch}$.

3.8 Nickel binding behaviour of apoACS and the apoACS-AcsF$_{Ch}$ complex

3.8.1 Nickel titration by ITC

To characterize the nickel binding behaviour of apoACS and the apoACS-AcsF$_{Ch}$ complex, ITC experiments were carried out. The binding isotherms are shown in Figure 3.11. Analyzing these isotherms is complicated, for several reasons: 1) ACS has two nickel binding sites, Ni_p and Ni_d, which are expected to have different binding constants, as nickel should bind more strongly to Ni_d than to Ni_p. 2) Other metals could have been bound to ACS prior to the titration, as ACS is known for its propensity to bind the "wrong" metal. 3) The titration curve of the apoACS-AcsF$_{Ch}$ complex does not level off at a stable plateau value, indicating that the titration had not completely finished. Increasing the nickel concentration in the syringe, however, is not feasible, as concentrations above 1 mM $NiCl_2$ can lead to protein precipitation. 4) Due to the low Wiseman "c" value of these ITC experiments, the inflection point of the curve, which is equal to the binding stoichiometry, is not clearly defined.

3.8.2 Metal analysis of Ni-titrated samples by ICP-OES

To determine the binding stoichiometry precisely, the samples from the ITC experiment were also analyzed by ICP-OES (Table 3.2): apoACS bound less than one nickel ion, i.e. 0.43 nickel ions per 4 iron atoms. This agrees with the fact that ACS cannot be easily activated by a simple nickel titration at room temperature. The apoACS-AcsF$_{Ch}$ complex, on the other hand, bound nickel almost stoichiometrically, i.e. 1.85 nickel ions per 4 iron atoms. On its own, apoACS bound nickel only substoichiometrically, while AcsF$_{Ch}$ did not bind nickel at all. There are three possible explanations for this result: 1) the binding affinity of nickel to ACS is increased in the complex, 2) due to conformational changes, a new nickel binding site is created in AcsF$_{Ch}$, or 3) a new nickel binding site is created at the interface of the two proteins.

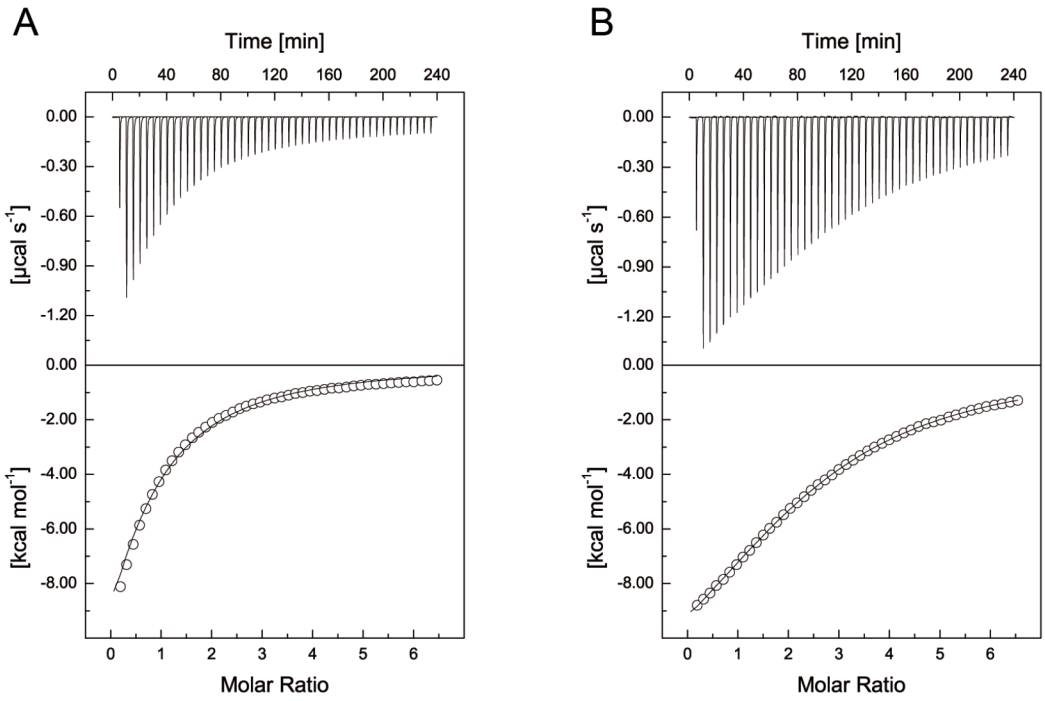


Figure 3.11: Isothermal titration calorimetry of apoACS and apoACS-AcsF$_{Ch}$ with $NiCl_2$. Top panels show the ITC raw data and bottom panels show the heat change of injections, which are normalized with respect to the moles of $NiCl_2$ added. A) ApoACS with $NiCl_2$. The data was fitted to a one-site model, which yielded a dissociation constant K_D = 36 µM and a stoichiometry N = 0.52 ± 0.11 binding sites. B) ApoACS-AcsF$_{Ch}$ complex with $NiCl_2$. The data was fitted to a one-site model, which yielded a dissociation constant K_D = 38 µM and a stoichiometry N = 2.52 ± 0.01 binding sites.

Table 3.2: Metal analysis of different ACS samples by ICP-OES. Values show the number of metal ions per molecule of ACS. The metal content measurements were replicated three times for all ACS samples and twice for the ACS-AcsF$_{Ch}$ complex sample. n.d. stands for not detected.

	Ni	**Zn**	**Fe**	**Ni per 4 Fe**
apoACS	n.d.	1.33 ± 0.18	4.83 ± 0.11	0
	n.d.	0.73 ± 0.11	3.25 ± 0.16	0
ACS after ITC	0.54 ± 0.14	n.d.	4.96 ± 0.22	0.43
ACS-AcsF$_{Ch}$ after ITC	2.26 ± 0.05	n.d.	4.88 ± 0.09	1.85

3.9 Conversion of the inactive Ni-ACS-AcsF$_{Ch}$ complex into active ACS

The metal analysis by ICP-OES showed that the ACS-AcsF$_{Ch}$ complex binds two nickel ions. When the ACS activity of this sample was measured, however, no activity was detected. Even though two nickel ions are bound to the complex, they are not able to catalyze the formation of acetyl-CoA. To test whether Mg-ATP can convert this inactive species to active ACS, the sample was incubated with 10 mM Mg-ATP at 60 °C for one hour, and the activity was measured again. After treatment with Mg-ATP, the sample was active, showing that Mg-ATP hydrolysis is involved in the final activation step, in which the inactive Ni-ACS-AcsF$_{Ch}$ complex is converted into active ACS (Fig. 3.12).

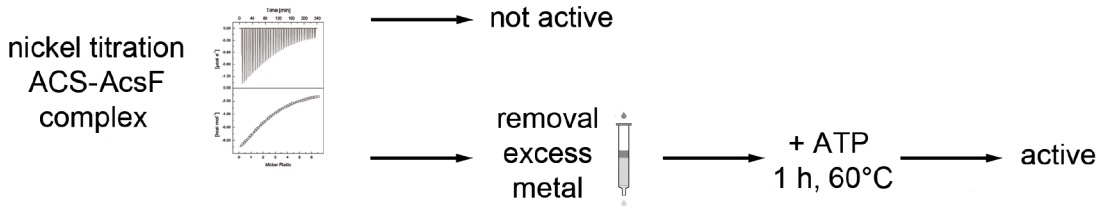


Figure 3.12: Stepwise activation of ACS. The sample from the ITC experiment is not active even though two nickel ions are bound to the complex. When Mg-ATP is added, however, ACS becomes active.

3.10 Detailed sequence analysis of AcsF- and CooC-type proteins

The biochemical experiments carried out as part of this project clearly demonstrate that AcsF$_{Ch}$ is a maturation factor of ACS. Furthermore, the experiments have revealed that AcsF$_{Ch}$ cannot be functionally replaced by the CooC-type proteins CooC1$_{Ch}$ or CooC3$_{Ch}$. AcsF- and CooC-type proteins belong to the same family of proteins as defined by the Interpro database, but can be divided into different subgroups (Fig. 3.13). To analyze the sequential differences of these two subgroups in greater detail, a sequence bundle alignment was carried out. This alignment visualizes whether certain residues or motifs are unique to a particular subgroup of proteins.

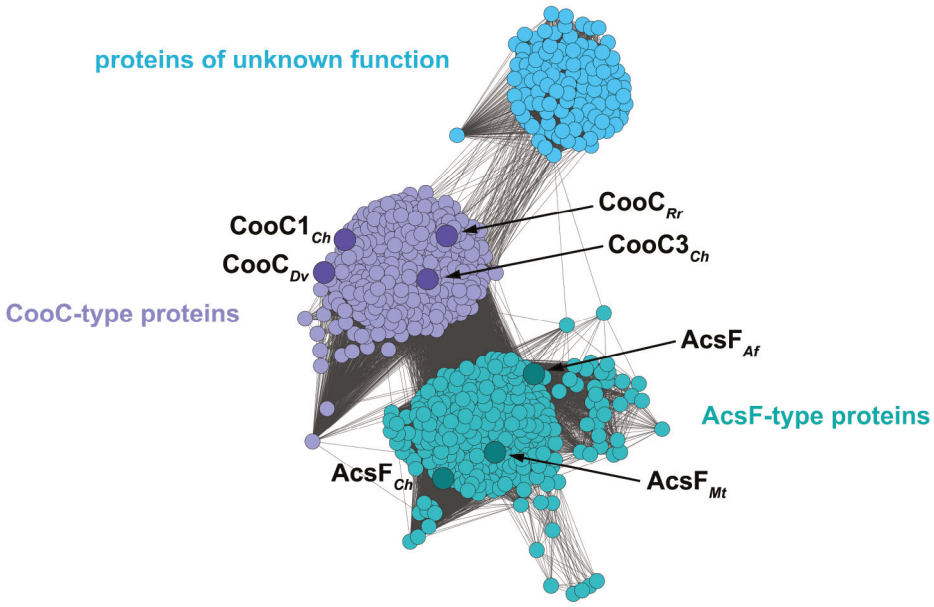


Figure 3.13: Sequence similarity network of the family IPRO14433. The three different subgroups – proteins of unknown function, CooC-type and AcsF-type proteins – are shown in light blue, violet and turquoise, respectively. Each node represents an amino acid sequence. Sequences from the genomes of *C. hydrogenoformans* ($CooC1_{Ch}$, $AcsF_{Ch}$, $CooC3_{Ch}$), *M. thermoacetica* ($AcsF_{Mt}$), *R. rubrum* ($CooC_{Rr}$) and *D. vulgaris* ($CooC_{Dv}$) are shown as larger nodes. The distribution of sequences from archaeal and bacterial proteins is shown in Figure A.7 in the Appendices.

The Walker A motif is highly conserved in both groups (Fig. A.8 in the Appendices). It has the consensus sequence GKGGXGKT. In CooC-type proteins "X" is usually a valine residue (88 %), whereas in AcsF-type proteins "X" is either a valine (43 %) or a threonine residue (53 %).

The switch I motif is characterized by two aspartate residues that are completely conserved in both subgroups (Fig. 3.14). Other residues that are highly conserved in both subgroups are leucine (94 %) in site 124, as well as leucine (88 %) and glycine (90 %) in site 128 and 129, respectively. The residue that is positioned between the two aspartate residues is conserved in the AcsF subgroup (alanine, 97 %), but not in the CooC supgroup.

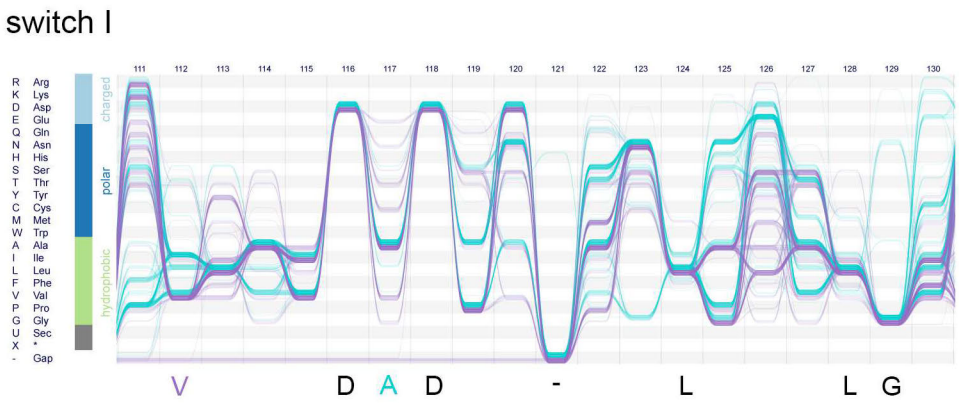


Figure 3.14: Sequence bundle alignment of the switch I motif. The black single-letter amino acid code is depicted below the alignment, if the residue is conserved in more than 85 % of all sequences. The code is depicted in violett and turquois, if the residue is conserved in more than 85 % of sequences of the CooC- and AcsF-type subgroup, respectively. Numbers in the top refer to the number of the site in the alignment.

The switch II motif has the consensus sequence DXEAGXEHXXR (Fig. A.9 in the Appendices). The sequences of the two subgroups display a high degree of similarity and there are no obvious features that are specific to only one of the two subgroups.

The cysteine residues of the CXC motif are completely conserved within both subgroups. Strikingly, the alignment reveals that the residue positioned between the two cysteine

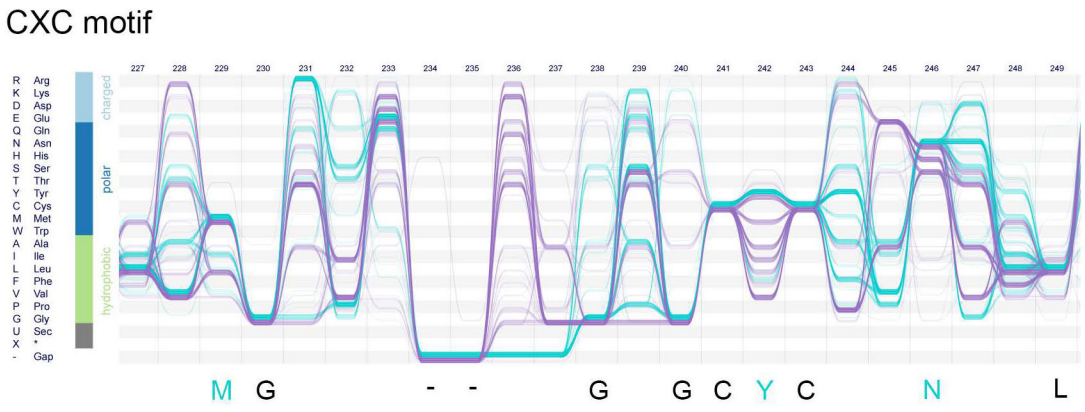


Figure 3.15: Sequence bundle alignment of the CXC motif. The colour coding is the same as in Fig. 3.14

residues is highly conserved within the AcsF subgroup, but not within the CooC subgroup (Fig. 3.15). In the AcsF subgroup, an aromatic residue, either tyrosine (96 %) or phenylalanine (4 %), is present at this site. Overall, sequences of the AcsF subgroup seem to be more conserved in close proximity to the CXC motif than sequences of the CooC subgroup (Fig. 3.15). An intact CXC motif is crucial to the functioning of both AcsF- and CooC-type proteins (see sections 3.6, 3.7 and [64]). Yet biochemical experiments have shown that each subgroup exhibits a different metal binding behaviour: CooC1$_{Ch}$ binds nickel [64], while AcsF$_{Ch}$ does not (see section 3.3). Whether this is due to the conformation of the CXC motif itself, or to a different arrangement of the dimer, remains to be seen. As described in section 3.8, the CXC motif might be part of a new metal binding site in the ACS-AcsF complex. The sequence alignment strongly suggests a requirement for an aromatic residue between the two cysteines. The functional role is not yet known, but it could stabilize interactions between the two proteins.

The most prominent deviation between the AcsF- and CooC-type proteins was found from sites 138 to 148 (Fig. 3.16). The residue in site 140 differs between the subgroups,

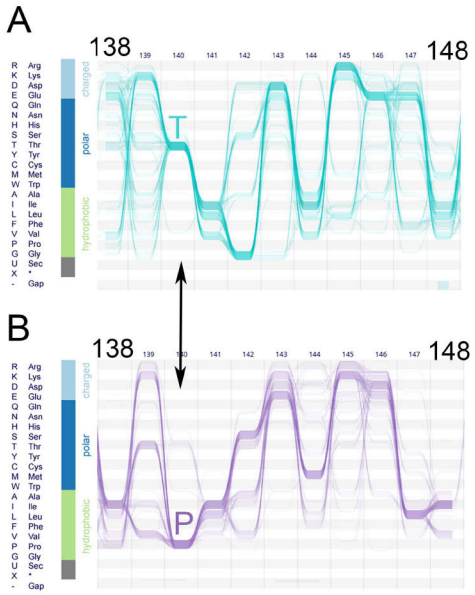


Figure 3.16: Sequential differences at site 140. A) Sequence bundle alignment of AcsF proteins from site 138 - 148. B) Sequence bundle alignment of CooC proteins from site 138 - 148.

but is highly conserved within each one: in 93 % of AcsF-type proteins a threonine residue is located at this site, while in 92 % of CooC-type proteins a proline residue is found.

3.11 Protein crystallization and structure determination

3.11.1 Crystallization and structure determination of $AcsF_{Ch}$

Crystals of $AcsF_{Ch}$ appeared in a condition which contained the ACS-$AcsF_{Ch}$ complex (45 mg/mL) in 25 mM Tris pH 8.0, mixed with the reservoir solution of 0.2 M lithium sulfate, 0.1 M phosphate citrate pH 4.2 and 20 % PEG 1000. The crystallization plate was set up in the glove box, as described in section 2.14, but it was taken out after approximately six months because no crystals had appeared. It took at least another six months until crystals of $AcsF_{Ch}$ appeared. The crystals had a pillow-like shape, as shown in Figure 3.17. Since the ACS-$AcsF_{Ch}$ complex – and not $AcsF_{Ch}$ alone – was originally present in the crystallization condition, the complex had obviously dissociated during this twelve-month period. Moreover, the crystal structure (see section 3.12.1) revealed that $AcsF_{Ch}$ had become oxidized during this time. Taken together, these results may explain why it took so long for crystals to appear.

Figure 3.17: Protein crystals of $AcsF_{Ch}$ in 0.2 M lithium sulfate, 0.1 M phosphate citrate pH 4.2 and 20 % PEG 1000. The picture on the right shows a magnification of the top right corner of the crystallization drop.

The crystals, which diffracted up to 2.7 Å, belonged to the tetragonal spacegroup $P4_12_12$, with unit cell dimensions a = b = 63.9 Å and c = 228.4 Å. Molecular replacement with the search model "short-CooC1" (Fig. A.2 in the Appendices) gave a solution with a translation function Z-score of 5.8 and a log likelihood gain of 40, indicating that a correct solution would be unlikely. Nonetheless, model building with AutoBuild was successful. The model consisted of residues 12 - 99, 115 - 134, 184 - 196, 209 - 252 and five additional fragments, which had a sequence identity of 64.5 %. At this stage the R_{work} and R_{free} were 30.5 % and 34.6 %, respectively. The model was completed and refined to an R_{work} of 26.1 % and an R_{free} of 28.7 %. The crystallographic data table, which contains the data collection and refinement statistics, is shown in Table A.2 in the Appendices.

3.11.2 Crystallization and structure determination of $AcsF_{Af}$

The AcsF protein from *A. fulgidus*, $AcsF_{Af}$, crystallized much more readily than $AcsF_{Ch}$. Crystals appeared in several conditions containing salt as precipitant. The structure of dimeric $AcsF_{Af}$ was determined from crystals which grew in a condition containing 0.8 M lithium sulfate monohydrate and 0.1 M Tris-HCl pH 8.5. These crystals, which diffracted up to 2.0 Å, belonged to the orthorhombic spacegroup $C222_1$ with unit cell dimensions a = 110.3 Å, b = 117.1 Å and c = 55.8 Å. Molecular replacement with the

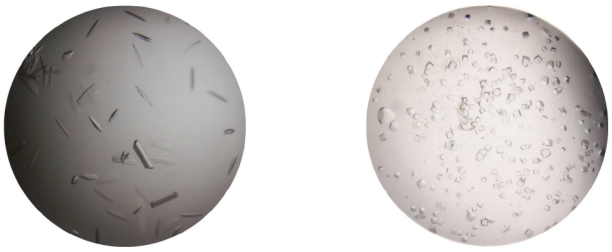


Figure 3.18: Protein crystals of $AcsF_{Af}$. The picture on the left shows crystals of dimeric $AcsF_{Af}$, while the picture on the right shows crystals of monomeric $AcsF_{Af}$.

manually prepared model "short-AcsF" (Fig. A.3 in the Appendices) yielded a translation function Z-score of 6.5 and a log likelihood gain of 179. Initial building with Autobuild gave a model containing the residues 1 - 66, 77 - 103, 106 - 135, 148 - 175 and 205 - 258, plus two additional fragments with a sequence identity of 72 %. The R_{work} and R_{free} of this model were 27.2 % and 32.1 %, respectively. The model was completed and improved to an R_{work} and R_{free} of 19.9 % and 22.5 %, respectively (Table A.3 in the Appendices).

The structure of monomeric $AcsF_{Af}$ was determined from crystals grown in a condition containing 0.5 M potassium thiocyanate and 0.1 M sodium acetate trihydrate pH 4.6. The crystals diffracted up to 1.7 Å, but unfortunately the diffraction pattern showed rings that originated from the presence of crystalline ice (Fig. A.11 in the Appendices). Before scaling, the data in the resolution range of the ice rings was excluded from the data set. Molecular replacement was carried out with $AcsF_{Af}$ as the search model, as at that time the structure of dimeric $AcsF_{Af}$ was already available. The model was refined to an R_{work} and R_{free} of 19.5 % and 23.9 %, respectively. The exclusion of the ice rings leads to a low completeness of the dataset (Table A.3 in the Appendices).

3.11.3 Structure determination of $CooC3_{Ch}$

The structure was solved by molecular replacement using the search model "short-CooC1" (Fig. A.2 in the Appendices). Initially, with only one molecule placed in the

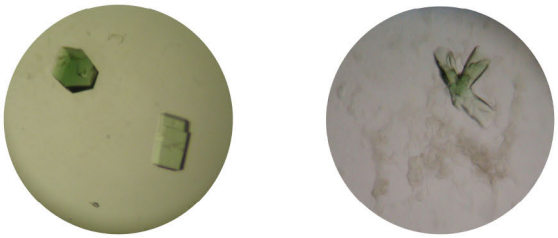


Figure 3.19: Green-coloured $CooC3_{Ch}$ crystals. These pictures were kindly provided by Dr. Jae-Hun Jeoung.

asymmetric unit, the structure refinement by PHENIX Refine yielded an R_{free} of 46.6 %. After several cycles of model building with COOT and refinement with PHENIX Refine, a new molecular replacement was performed with two molecules in the asymmetric unit, which decreased the R_{free} to 40.4 %. This model was subsequently further refined to an R_{work} and R_{free} of 21.4 % and 25.4 %, respectively. The structure was refined without restraints for the metal binding site.

3.12 Crystal structures

3.12.1 Overall structure of AcsF$_{Ch}$

In the crystal structure, AcsF$_{Ch}$ is present in a dimeric assembly, which is made up of two symmetry-related monomers. One monomer consists of seven parallel β-strands, which are surrounded by ten helices (Fig. 3.20). The β-sheet, which forms the central core of the protein, is characteristic of the SIMIBI class of proteins.

The deviant Walker A motif is located between the first β-strand β1 and the first helix α1. In the crystal structure, no nucleotide is bound to the Walker A motif. However, a phosphate molecule, most likely derived from the crystallization buffer, was modelled into the nucleotide binding site. The phosphate molecule occupies the same position as the β-phosphate of ADP in the crystal structure of CooC1$_{Ch}$ [76], and forms hydrogen bonds to the backbone of Gly12, Thr13 and Gly14 of the deviant Walker A motif (Fig. A.12 in the Appendices).

The switch I motif (VDADAN) is positioned between the second β-strand β2 and the second helix α2. As with other structures of the SIMIBI family, it is in close proximity to the nucleotide binding site. Aspartate 40 forms a hydrogen bond to a water molecule, which in turn is in hydrogen bonding distance to the phosphate molecule. The switch II motif (DNEAGMEHLSRRT) is located between β-strands β4 and β5. The CXC motif is part of a loop located before helix α5. The position of these motifs relative to each other is shown in Fig. 3.21.

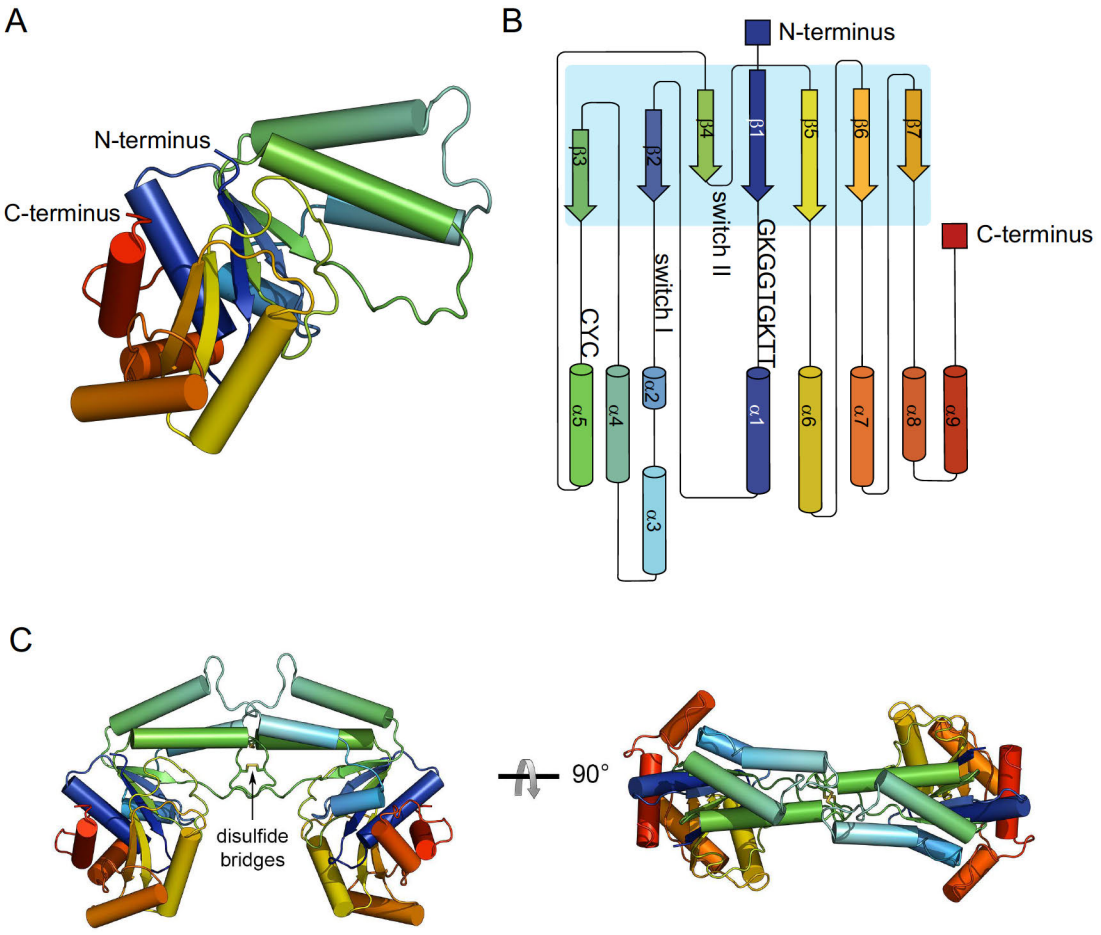


Figure 3.20: A) Crystal structure of $AcsF_{Ch}$ with a rainbow coloring scheme from the N-terminus (blue) to the C-terminus (red). B) Topology diagram of $AcsF_{Ch}$ with the same coloring scheme as in the structure. α-helices are represented as cylinders and β-strands as arrows. C) Crystal structure of dimeric $AcsF_{Ch}$. The cysteine residues that form the disulfide bridges are shown as sticks.

The cysteine residues form two disulfide bridges across the dimer interface (Fig. 3.20 C). Each disulfide bond is formed between equivalent residues from different monomers, i.e. Cys108 of one monomer is linked to Cys108 of the other, while Cys110 is linked to Cys110. As described in section 2.14, there was a very long delay between setting up the crystallization plate and the appearance of the crystals, during which the cysteine residues became oxidized. The oxidized protein is most likely not functional, since the disulfide bonds rigidify the protein structure and the two monomers seem unable to move relative to each other. However, the oxidation of the protein may have been beneficial to the crystallization of $AcsF_{Ch}$, as no crystals were obtained from reduced $AcsF_{Ch}$, which is most likely more flexible than the oxidized protein.

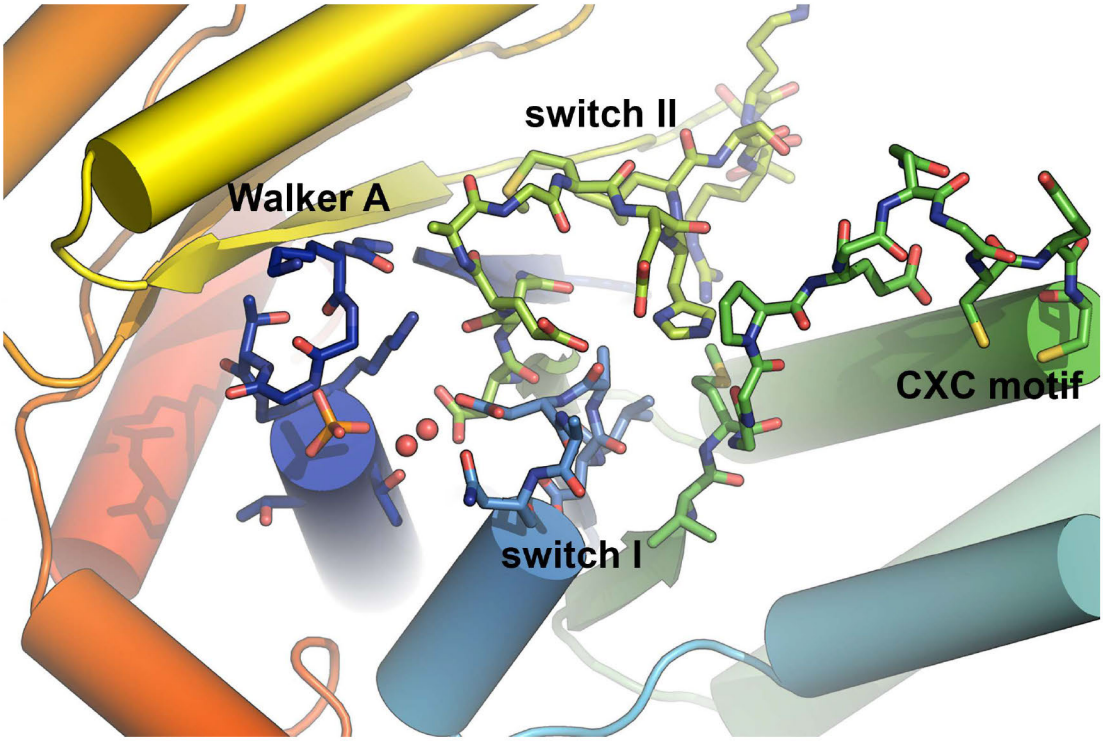


Figure 3.21: Positions of the structural motifs relative to each other. Residues of the Walker A motif are depicted as blue sticks, residues from the switch I and switch II motif are shown as light blue and light green sticks, respectively and residues of the CXC motif are represented by green sticks. One can easily imagine how structural rearrangements in one motif affect the conformation of the other.

An analysis of the temperature factors of the AcsF$_{Ch}$ crystal structure shows that helix α6 and the loop immediately following the switch II motif are the most flexible parts of the oxidized protein (Fig. A.13 in the Appendices).

The dimer interface is predominantly formed by residues of the loop between helices α3 and α4 as well as residues of the loop between β-strand β3 and helix α5 (Fig. 3.20 C). The buried surface interface area of the dimer is 1081 Å^2, which is equal to 4.4 % of the total surface area of the dimeric structure.

A DALI structural similarity search [116] shows that AcsF$_{Ch}$ has the highest structural similarity to CooC1$_{Ch}$ (Z-score = 22.0) followed by MinD (Z-score = 20.5) and ParA (Z-score = 18.9). A list with further similar structures is shown in Table A.5.

3.12.2 Overall structure of AcsF$_{Af}$

The overall structure of AcsF$_{Af}$ is very similar to AcsF$_{Ch}$. The crystal structure is comprised of a central β-sheet surrounded by 10 helices (Fig. 3.22). The Walker A motif is located between β-strand β1 and helix α1, the switch I motif (DADPDSN) between β-strand β2 and helix α2, the switch II motif (DTEAG) between β-strand β5 and helix α6 and the CXC motif is positioned before helix α5. There was no electron density present, which could account for a nucleotide binding in the P-loop. However, a phosphate molecule, which derived from the crystallization buffer, was present in the same position as in AcsF$_{Ch}$. A zinc atom is bound to the metal binding site. It is coordinated by Cys108 and Cys110 of opposite monomers in a tetrahedral coordination geometry (Fig. 3.22 C and Fig. 3.23).

As described in section 3.11.2 crystal structures of both monomeric and dimeric assemblies of AcsF$_{Af}$ have been solved. The two structures are very similar. The only diffference is that the loop containing the metal binding residue is very flexible in the monomeric structure as indicated by a poorly defined electron density and high temperature factors (Fig. A.14 in the Appendices).

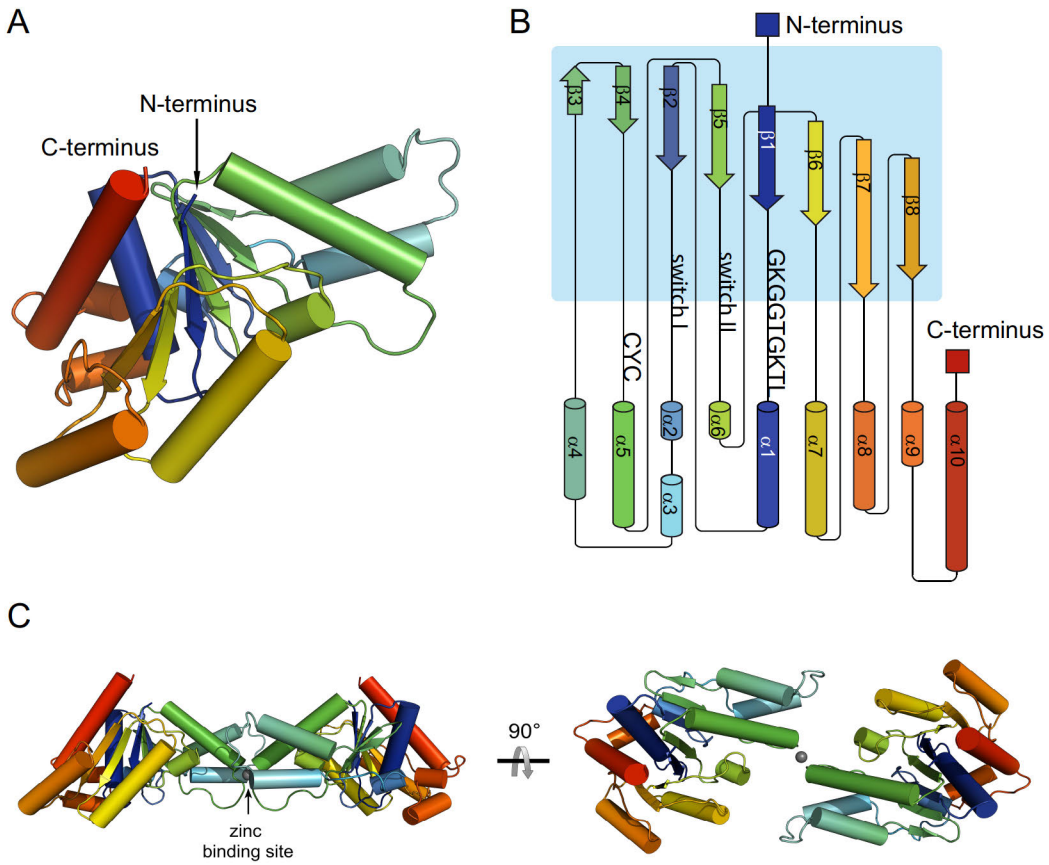


Figure 3.22: A) Crystal structure AcsF$_{Af}$ with a rainbow coloring scheme from the N-terminus (blue) to the C-terminus (red). B) Topology diagram of AcsF$_{Af}$ with the same coloring scheme as in the structure. α-helices are represented as cylinders and β-strands as arrows. C) Structure of dimeric AcsF$_{Af}$. Zinc is depicted as grey sphere.

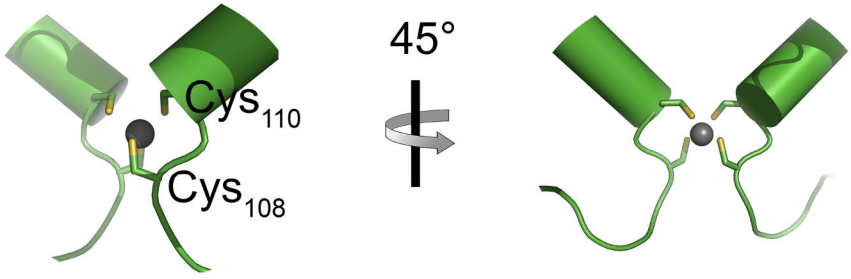


Figure 3.23: Zinc binding site of AcsF$_{Af}$. Zinc is depicted as grey sphere and cysteine residues are shown as sticks.

3.12.3 Overall structure of CooC3$_{Ch}$

The overall structure and fold of CooC3$_{Ch}$ is very similar to that of CooC1$_{Ch}$, which is described in detail in reference [76]. In the crystal structure of CooC3$_{Ch}$ the nucleotide binding site is empty and three nickel ions are bound to one dimer.

3.12.4 Nickel binding properties of CooC3$_{Ch}$

The crystals of CooC3$_{Ch}$ were green-coloured (Fig. 3.19), which indicated the presence of nickel in the crystal. The structure refinement revealed a stoichiometry of three nickel ions per dimer of CooC3$_{Ch}$. One of the three nickel ions is bound to the interface of the dimer, which is formed by monomers A and B' from different asymmetric units (Fig. 3.24). The nickel ion is coordinated in a square planar coordination geometry by the two cysteine thiolates (Cys123 and Cys125) of each monomer. The other two nickel ions are located adjacent to the central nickel ions, and occupy equivalent positions in monomers A and B', respectively. They are coordinated by the deprotonated amides of the protein backbone (Ala124 and Cys125) and by the cysteine thiolates (Cys123 and Cys125). The three nickel ions lie on a straight line, but the planes which are formed by the square planar coordinations have a stair-like arrangement with a dihedral angle of 106.0 ° and 102.7 °, respectively.

The stoichiometry of three nickel ions binding to one dimer was rather unexpected given its similarity to CooC1$_{Ch}$. Metal binding studies showed that in solution CooC1$_{Ch}$ can only bind one nickel ion per dimer at the CXC motif [64]. The detailed metal binding properties of CooC3$_{Ch}$ in solution have not yet been determined and further studies are required to show whether CooC3$_{Ch}$ is also able to bind three nickel ions in solution.

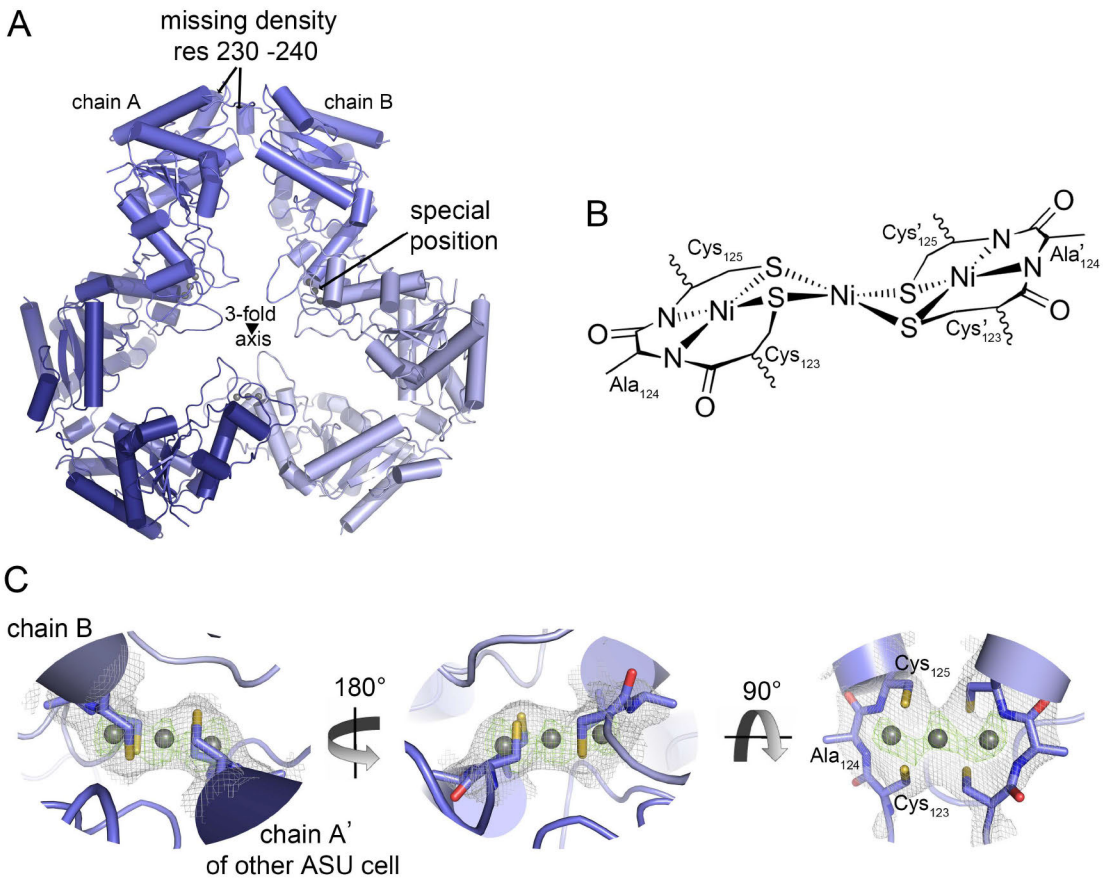


Figure 3.24: Nickel binding site of CooC3$_{Ch}$. A) Structure of three CooC3$_{Ch}$ dimers. Monomers, which are located in the same asymmetric unit are shown in the same tone of blue. One dimer consists of two monomers A and B' of different asymmetric units. Nickel ions are shown as grey spheres. One of the three nickel ions is bound to the interface of two monomers and is located on a crystallographic special position. B) Schematic representation of the metal binding site. C) Metal binding site from three different perspectives. The 2mFo-DFc electron density map of the metal binding site, contoured at 1.0 σ, is shown in grey and the anomalous difference map, contoured at 5.0 σ, is shown in green.

The nickel binding site formed by the nitrogen atoms of the backbone and the cysteine residues is analogous to the Ni$_d$ site in ACS. The only difference is that the Ni$_d$ site is formed by a Cys-Gly-Cys motif while the nickel binding site in CooC3$_{Ch}$ is formed by a Cys-Ala-Cys motif. In CooC3$_{Ch}$ this motif is located after a flexible loop with weak electron density before an α-helix, whereas in ACS this structural motif is part of a loop.

Apart from the biological metal complexes, inorganic chemists have synthesized nickel complexes which are remarkably similar to the metal binding site as found in CooC3$_{Ch}$. These have been referred to as tri-nuclear "stair-step" complexes [117, 118].

3.13 Structural comparison of AcsF and CooC proteins

To compare the overall structural differences between the two subgroups, the structures of the AcsF proteins AcsF$_{Ch}$ and AcsF$_{Af}$ were superimposed onto the structures of the CooC proteins CooC3$_{Ch}$ and CooC1$_{Ch}$. The alignment was restricted to the central β-sheet, which is highly conserved in all four structures, thereby illuminating the structural differences between the remaining elements.

The main structural variation is present between the switch I motif and the CXC motif (Fig. 3.25). The secondary structure elements are similar, but the orientation relative to the rest of the structure differs between the two subgroups. In the CooC proteins CooC1$_{Ch}$ and CooC3$_{Ch}$ the position of helices α3 and α4 is rotated by approximately 60 ° compared to the position of these helices in the AcsF type proteins AcsF$_{Ch}$ and AcsF$_{Af}$. Furthermore, the loop connecting these helices is positioned above the structure in AcsF-type proteins, but below the structure in CooC-type proteins. Of all four proteins AcsF$_{Ch}$ is the only protein that does not contain a β-strand that would be equivalent to the β-strand β3 of AcsF$_{Af}$.

The sequence bundle alignment revealed sequential differences from sites 138 to 148. These residues are positioned between the switch I and the CXC motif. Thus they are also located in the region which differs structurally between the AcsF and CooC proteins. A closer inspection reveals that the conserved proline residue in CooC proteins (site 140) is located at the start of a short 3_{10} helix, whereas the equivalent residue in AcsF proteins, a conserved threonine residue, is the first residue of an α-helix (Fig. 3.26).

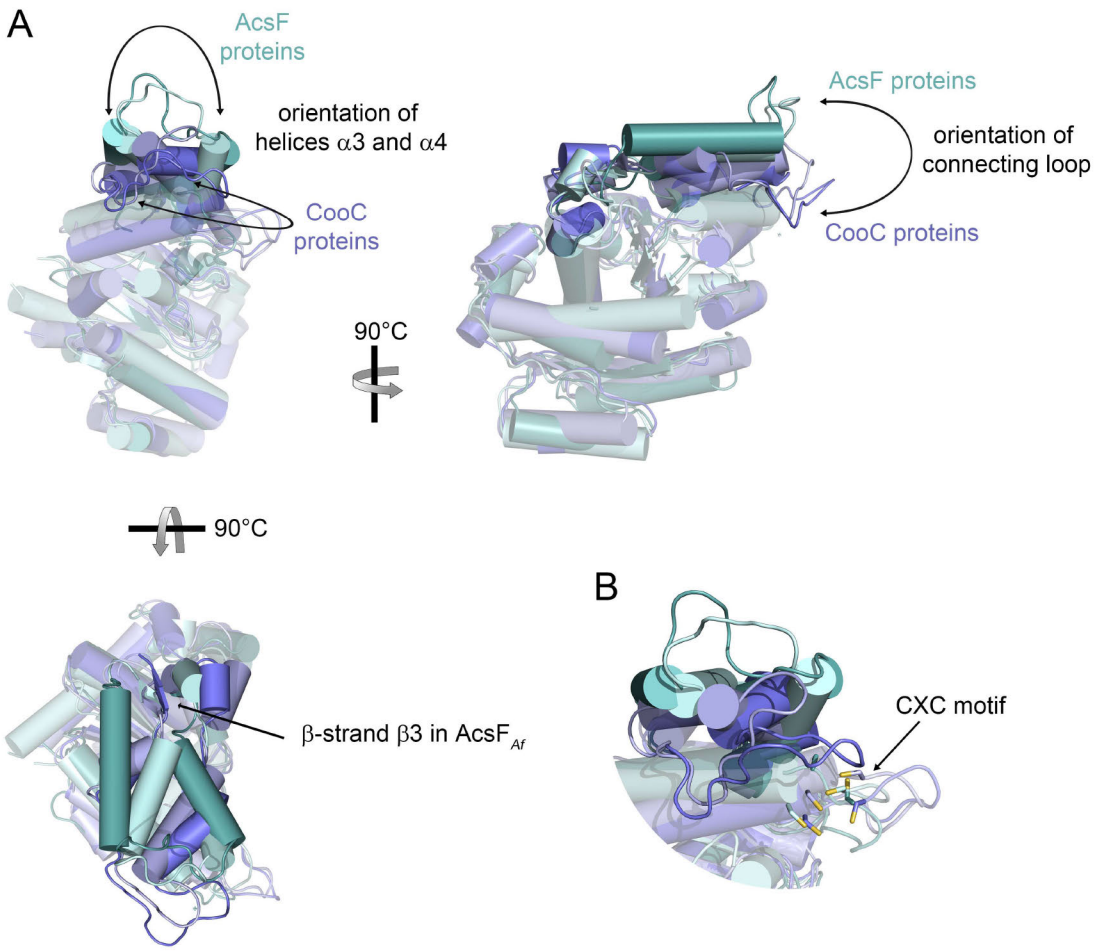


Figure 3.25: A) Structural comparison of AcsF- and CooC-type proteins. Structures of the AcsF proteins, AcsF$_{Ch}$ and AcsF$_{Af}$, are shown in dark and light turquoise, respectively. Structures of the CooC proteins, CooC1$_{Ch}$ and CooC3$_{Ch}$, are shown in light and dark violet, respectively. The conserved parts of the structure are shown as transparent cartoon representations. B) Location of the CXC motif in the four structures. The side chain of the cysteine residues is shown as sticks and sulfur atoms in yellow.

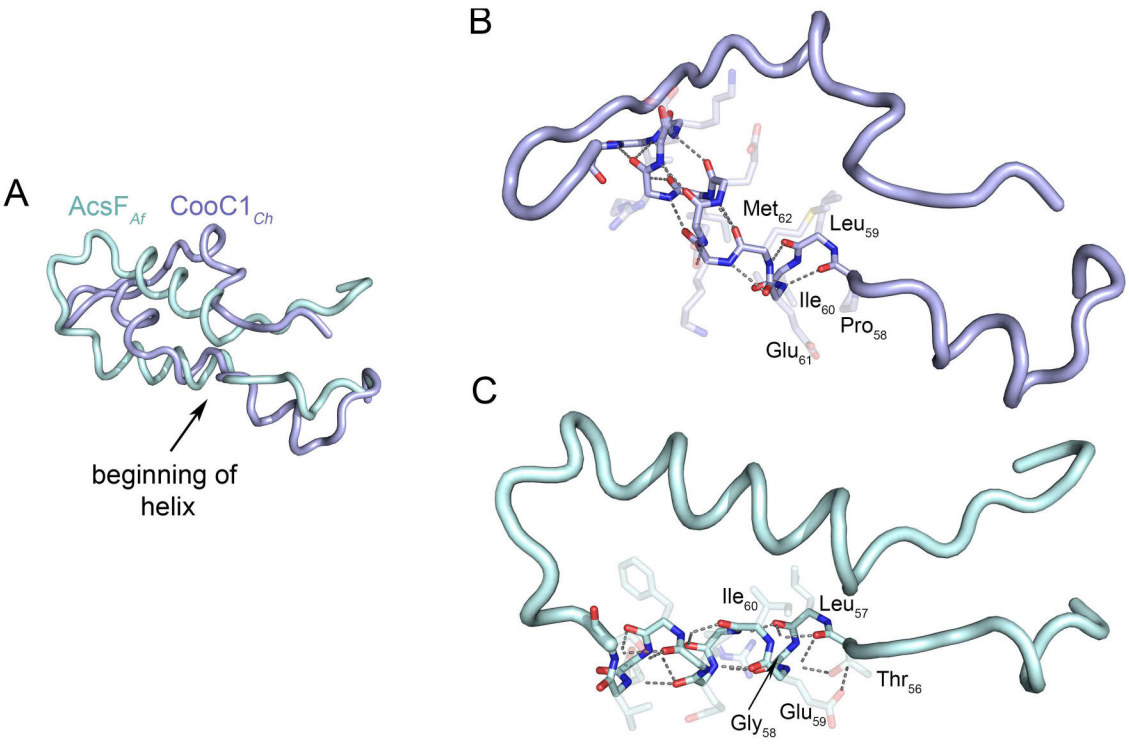


Figure 3.26: Detailed comparison of helix α3 of AcsF$_{Af}$ and the equivalent positions in CooC1$_{Ch}$. A) Superimposed structures of AcsF$_{Af}$ and CooC1$_{Ch}$ of residues 47 - 95 and 45 - 93, respectively. AcsF$_{Af}$ and CooC1$_{Ch}$ are shown in light turquoise and light violet, respectively. B) Close up view of the CooC1$_{Ch}$ structure with residues 58 - 70 shown as sticks. The backbone is shown as opaque sticks, and side chains are shown as transparent sticks. Hydrogen bonds are shown in grey. The proline residue, which is a conserved feature of the CooC proteins, is positioned at the start of a short 3_{10} helix which consists of five residues (PLIEM) that form two i+3 → i hydrogen bonds. The first hydrogen bond is formed between proline and glutamate and the second one is formed between leucine and methionine. The methionine residue is at the same time the first residue of the α-helix which consist of residues 62 - 70 (MKDEIREKT). C) Detailed representation of helix α3 in AcsF$_{Af}$. The helix is an α-helix comprised of 11 residues. The first residue of the helix is threonine, which is a conserved feature of the AcsF group. The backbone of the helix residues is shown as opaque sticks, and side chains are shown as transparent sticks.

Surprisingly, the CXC motif is located in a similar position in all four proteins (Fig. 3.25 B). In the structure of AcsF$_{Ch}$, the cysteine residues point inside, while in all other structures they point outside. A structural alignment was carried out in which only the cysteine residues were superimposed on top of each other (Fig. 3.27). The conformation of the protein backbone is identical in AcsF$_{Af}$ and CooC1$_{Ch}$, while it differs within each subgroup. This is also in line with the fact that AcsF$_{Af}$ and CooC1$_{Ch}$ both bind zinc, while CooC3$_{Ch}$ binds nickel and AcsF$_{Ch}$ forms disulfide bridges in the crystal structure. Taken together, this shows that the CXC motif is highly flexible and no clear distinction between AcsF and CooC proteins is apparent. This supports the theory that the conserved aromatic residue of the AcsF proteins might be important for its interaction with ACS.

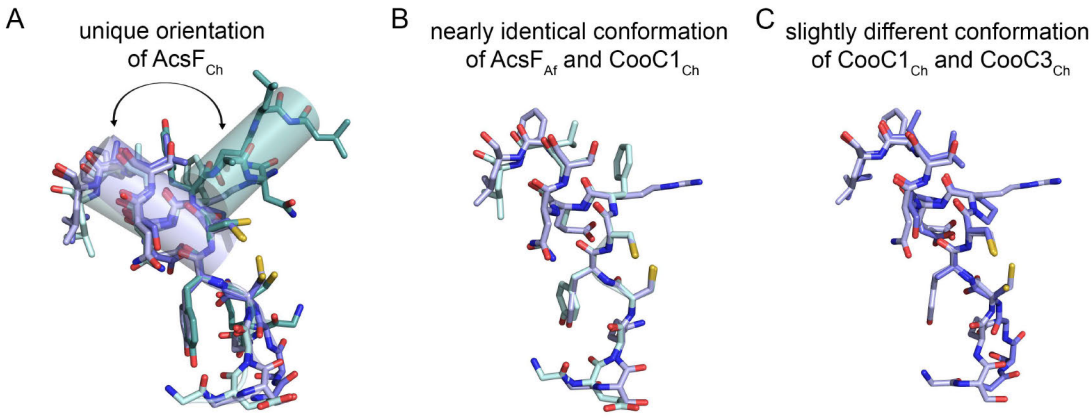


Figure 3.27: Superimposed CXC motif of AcsF and CooC proteins. A) The CXC motif of AcsF$_{Ch}$ has a unique conformation, while B) AcsF$_{Af}$ and CooC1$_{Ch}$ adopt a nearly identical conformation. C) The conformation of CooC1$_{Ch}$ and CooC3$_{Ch}$ differs slightly.

It is important to note that certain conformational changes observed in protein crystals can arise from packing forces. For various reasons, however, this was unlikely the case. Firstly, structures from two proteins were available for each subgroup. Secondly, two of the four proteins crystallized in several different conditions. The crystals had different unit cell parameters and therefore differed in crystal packing, but the monomers from

different conditions were superimposable.

In addition to the visual inspection of the superimposed structures, a FATCAT (Flexible structure AlignmenT by Chaining Aligned fragment pairs allowing Twists) [119] was carried out (Table 3.3). The alignment agrees with the previous results and shows that overall the most similar structure to $AcsF_{Ch}$ is $AcsF_{Af}$. Of the two AcsF proteins, the archaeal protein $AcsF_{Af}$ is more similar to the CooC-type proteins than the bacterial protein $AcsF_{Ch}$.

Table 3.3: FATCAT structural alignment of AcsF and CooC proteins

protein 1	protein 2	equivalent positions	r.m.s.d.
$AcsF_{Ch}$	$AcsF_{Af}$	245	3.01
$AcsF_{Ch}$	$CooC1_{Ch}$	219	3.09
$AcsF_{Ch}$	$CooC3_{Ch}$	208	3.09
$AcsF_{Af}$	$CooC1_{Ch}$	237	3.08
$AcsF_{Af}$	$CooC3_{Ch}$	227	3.06
$CooC1_{Ch}$	$CooC3_{Ch}$	239	1.91

The CooC proteins $CooC1_{Ch}$ and $CooC3_{Ch}$ are more similar to each other than the AcsF proteins $AcsF_{Ch}$ and $AcsF_{Af}$. This was also reflected by the sequence similarity network (Fig A.7 in the Appendices), which shows a slight separation between archaeal and bacterial proteins within the AcsF subgroup. One might conclude that the AcsF-type proteins differ slightly between archaeal and bacterial organisms, since the β-subunit of the archaeal ACDS complex does not contain the N-terminal domain that is present in the bacterial enzyme. If the β-subunit is matured while already part of the ACDS complex, AcsF might also interact with other subunits of the ACDS complex. So far, however, this is only speculation, as the order of assembly and maturation of the ACDS complex is not known.

This also highlights why ACS and AcsF from *C. hydrogenoformans* are an attractive model system with which to study the interaction between those two proteins. ACS from *C. hydrogenoformans* has been shown to be present as a monofunctional enzyme

in vivo, meaning the interplay with other proteins need not be taken into account.

4 | Conclusion

AcsF is a maturation factor of ACS that catalyzes the nickel insertion into the A-cluster of ACS in an Mg-ATP dependent process. Biochemical experiments have shown that two molecules of AcsF$_{Ch}$ form a complex with one molecule of ACS. Nickel binding studies revealed that the apoACS-AcsF$_{Ch}$ complex binds nickel stoichiometrically (two nickel ions), while apoACS alone is only capable of substoichiometric nickel binding. Therefore, either a new metal binding site is created in the complex, or the affinity of the nickel binding sites in ACS is increased. The exact position of the nickel binding sites in the ACS-AcsF$_{Ch}$ complex is not yet known and further experiments are necessary to determine their location. The ACS-AcsF$_{Ch}$ complex containing two nickel ions is not active. To convert the inactive species to the active one, ATP hydrolysis by AcsF is required. Figure 4.1 shows a model that summarizes the proposed role of AcsF.

The biochemical experiments also clearly showed that AcsF proteins can not be functionally replaced by the related CooC proteins. Even though AcsF and CooC proteins are classified into the same family, a sequence similarity network revealed that AcsF and CooC proteins each form their own subgroups within this family. A comparison of the crystal structures of two AcsF proteins with the structures of two CooC proteins showed that the structural elements between the switch I and CXC motif are the least similar.

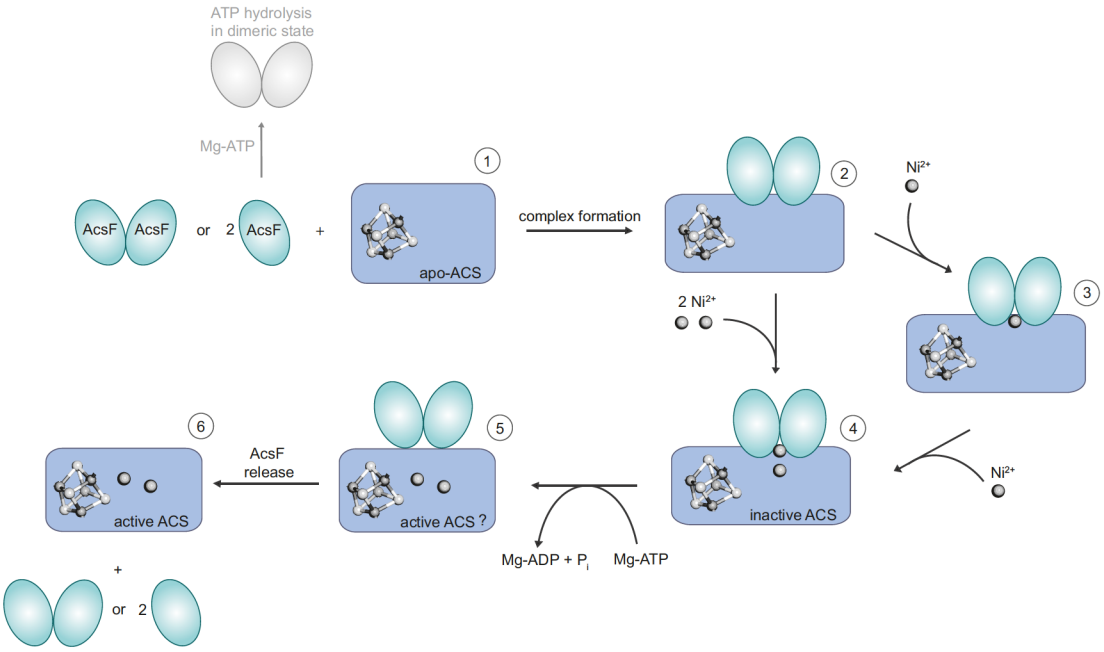


Figure 4.1: Preliminary model of ACS activation by $AcsF_{Ch}$. ApoACS (1), which contains the $[Fe_4S_4]$ cluster, but no nickel, can form a complex with either one dimer or two monomers of $AcsF_{Ch}$. $AcsF_{Ch}$ is an ATPase of the MinD-type family that dimerizes in the presence of Mg-ATP (grey depiction). The apoACS-$AcsF_{Ch}$ complex (2) is able to bind two nickel ions. The details of the nickel binding mechanism and the location of the nickel binding site are not yet known. The nickel ions may bind either sequentially or simultaneously. Even though two nickel ions are bound to the complex (4), the complex does not have acetyl-CoA formation activity. ATP hydrolysis by $AcsF_{Ch}$ is required to convert the inactive form of ACS (4) to an active form of ACS (5). Whether ACS is active within the complex, is not yet clear. In the last step, $AcsF_{Ch}$ dissociates from ACS.

A | Appendices

A.1 Appendices to Materials and Methods

```
CAT ATG ACA GTA ATC GCG GTG GCA GGC AAG GGT GGG ACC GGA AAA ACG CTG GTT TCC GCC CTC
CTC ATT AAT TTT ATT TCA GAG CAC ACT ACA AGC GTT CTG GCA GTT GAT GCC GAC CCA GAC AGC
AAC CTC CCA GAA GCT CTC GGT GTG GAG AAG CAG GTA AGA AAA ACT CTT GGC GAA ATC AGG GAG
CTC TTT CAG GTT AGC AGG GAT GAA ATG GGT TCA ATG AAC AAG GAG CAG TGG TTG GAA GGC AAG
ATT TAC GCT GAA GCA ATA TGC GAG TGC CCC AGA TAC GAC CTT CTG GTG ATG GGA AGG CCG GAA
GGA GAG GGG TGC TAC TGC TTT GCC AAT TCT CTG CTT CGA GGT GTG CTG AAG AGG CTG ATG AGA
CAC TAC GAG TAC ATA ATA ATC GAC ACC GAA GCT GGT CTG GAG CAC TTC AGC AGA AAA ACC ATC
GAC TCT GCT GAT TAC ATA ATA ATC GTA ACT GAC ATG TCG AGA AAG GGA CTG GCA ACT GCA AAG
AGG ATC AAA GAG CTT GCC AGT GAG CTC AAG CTG AAC TTC AAG AAG ATT TTC CTG ATT GCC AAC
AGA ATC GCT TCC GAG GAT GCA GAG AAA ACG ATT AGA GAG TTT GCG AAA GAG GAG GGT TTG GAG
CTT CTT GCC GTG CTG CCC TAC GAC AGC AGT GTA GCA GAA ATC GAC CTT AGA GGA GAA CCT GTA
AGC AAG ATA GAC AAG AAT TCT GAA GTC TAT AGG AAA ATG AAG GAT GTA GCT AAT TTA ATG CTC
AAT TTG TCA GCC AAA GCG AGG TGA GGA TCC
```

Figure A.1: Synthetic gene encoding for AcsF$_{Af}$. NdeI and BamHI restrictions sites are underlined.

Table A.1: PCR primers used for cloning. Restriction sites are underlined.

name	sequence
C2-fw	5'-GGA ATT CCA TAT GGC CTT TAA AAT TGC GGT TG-3'
C2-rv	5'-CAT GGT CTC GGA TCC TCA GAT AAT ACC CGC ATT TG -3'
C3-fw	5'-GGA ATT CCA TAT GCA GGT GTC TG-3'
C3-rv	5'-CCC AAG CTT CTA ACA ATT CTG CC-3'
PIPE-vec-fw	5'-GGA TCC GAA TTC GAG CTC CGT CG-3'
PIPE-vec-rv	5'-CAT ATG GCC CTG GAA ATA CAA GTT TTC GGT CG-3'
PIPE-C2-fw	5'-CTT GTA TTT CCA GGG CCA TAT GGC CTT TAA AAT TGC GGT TGC AG-3'
PIPE-C2-rv	5'-CGA CGG AGC TCG AAT TCG GAT CCC TAG ATA ATA CCC GCA TTT TGT AAA ATT TTC-3'

A

```
        10          20          30          40          50          60
IISIVSGKGG TGKTTVTANL SVALGEMGRK VLAVDGDLTM NLSLVLGVDD VNITLHDVLA

        70          80          90          100         110         120
GDAKLEDAIY MTQFENVYIL PGAVDWEHVI KADPRKLPEV IKSLKGKYDF ILIDCPAGLQ

        130         140         150         160         170         180
LRAMSAMLSG EEAILVTNPE ISCLTDTMKV GMVLKKAGLA ILGFILNRYG RSERDIPPEA

        190         200         210         220
AQDDVPLLAV IPEDPVIREG TLEGIPAVKY KPEAFIKLAE EVDKL
```

B

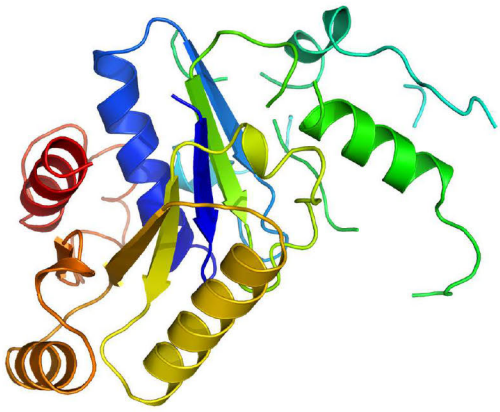


Figure A.2: A) Sequence of the search model “short-CooC1”. B) Cartoon representation of the search model “short-CooC1” with a rainbow coloring scheme from the N-terminus (blue) to the C-terminus (red). This model was prepared with PHENIX sculptur.

A

```
        10         20         30         40         50         60
FKIAVAGKGG TGKTTFSALV IKQLLMSGKR PILAVDADAN ANLNEALGDD VDLLVMGGPE

        70         80         90        100        110        120
GQGCYCYANN LLRQFILTLS NNYPYIVMDN EAGMEHLSRR TTDEVDVFFV ISDGSVRGIR

       130        140        150
SAGRIKQLID AIKEKYLVIT RKLIGVIPND ELVTE
```

B

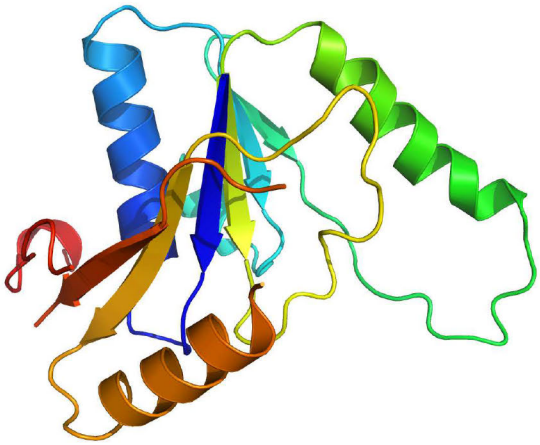


Figure A.3: A) Sequence of the search model “short-AcsF”. B) Cartoon representation of the search model “short-AcsF” with a rainbow coloring scheme from the N-terminus (blue) to the C-terminus (red). This model was prepared manually by truncating sections from the AcsF$_{Ch}$ structure.

A.2 Appendices to Results and Discussion

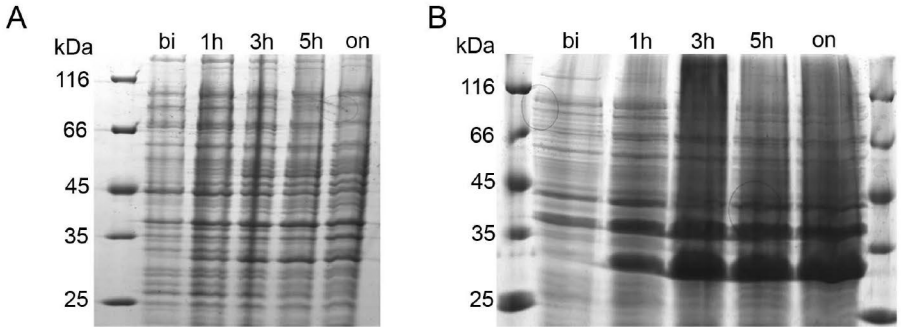


Figure A.4: Comparison of the expression level of the gene encoding for $AcsF_{Ch}$ in different vectors. A) Exemplary SDS-PAGE of samples taken at different time points from *E. coli* BL21(DE3) cultures, which contained the pASK-IBA17plus plasmid. bi stands for samples taken before induction, 1h, 3h and 5h stand for samples taken at 1, 3 and 5 hours after induction and on stands for samples taken from the culture which was grown overnight. The first lane contained a molecular weight marker with the indicated molecular masses. B) Exemplary SDS-PAGE of samples from *E. coli* BL21(DE3), which contained the pET28a plasmid. The labeling is the same as in part A of this Figure. A marker was present in the first and in the last lane. The molecular weight of $AcsF_{Ch}$ including the streptag, linker and TEV cleavage site is 30.6 kDa. Thus, the pronounced band between 25 kDa and 35 kDa can be assigned to $AcsF_{Ch}$. The expression level of the gene encoding for $AcsF_{Ch}$ is higher in BL21(DE3) cells containing the pET28a vector than in BL21(DE3) cells containing the pASK-IBA17plus vector.

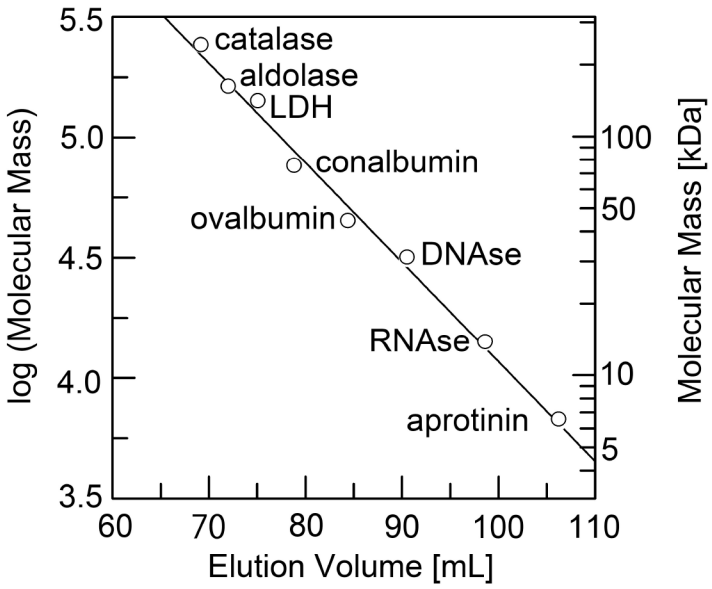


Figure A.5: Calibration curve of the superdex S200 Hiload 16/60 column. Molecular weight standards were aprotinin (6.5 kDa), RNase A (13.7 kDa), DNase I (31 kDa), ovalbumin (44 kDa), conalbumin (75 kDa), lactate dehydrogenase (140 kDa), aldolase (161 kDa) and catalase (240 kDa). The linear regression equation is: $y = -0.0413x + 8.201$.

```
AcsF_Ch          -----MAFKIAVAGKGGTGKTTFSALVIKQLLMSGKRPILAVDADANANLNEALGLQVEQ 55
CooC1_Ch         -------MKLAVAGKGGVGKTTVAAGLIKIMASDYDK-IYAVDGDPDSCLGQTLGLSIEE 52
CooC3_Ch         MQVSEKGLKIAVSGKGGVGKTTLSALLSHIFAREGKR-VLAVDADPDANLGTALGFPTEV 59
                        :*:**:****.****.:* : : :  . .: : ***.*.:: *. :**:  *

AcsF_Ch          AISDVINRLAKNMDPIPAG-----MTKDQYISFKIHETLSEG-----DDVDLLVMGGPE- 104
CooC1_Ch         AYAITPLIEMKDEIREKTG-----DGGLLILNPKVDGDLDKYGRYIDDKIFLIRMGEIKK 107
CooC3_Ch         LEKLTAISEDKELIKERTGAEPGTSGQYFILNPRVDDIPEKY-VVEHAGIRLMQMGKVVR 118
                     .     *:     :*           :. ::.   .:        : *: **

AcsF_Ch          -GQGCYCYANNLLRQFILTLSNN-YPYIVMDNEAGMEHLSRRTTDEVDVFFVISDGSVRG 162
CooC1_Ch         GGSQCYCRENSFLGSVVSALFLDKKEAVVMDMGAGIEHLTRGTAKAVDMMIAVIEPNLNS 167
CooC3_Ch         GGSGCACPESVLLKHLLRHLVLKADETVIVDMEAGLEHLGRGTAEGVDAFIVVVEPGKRS 178
                  *. * *  . :*  .:  *  .    :::*  **:*** * *:. ** ::.: : . ..

AcsF_Ch          IRSAGRIKQLIDSLDLKIKEKYLVITRIEEKDIPEVQEEIEKTGLKLIGVIPNDELVTEF 222
CooC1_Ch         IKTGLNIEKLAG--DLGIKKVRYVINKVRNIKEEKLIKKHLPE-DKILGIIPYNELFIEL 224
CooC3_Ch         FQTARAVVTLAR--DLGVTKIFAVANKVR-PGDEEIIRKELNF-LPILGFIPYDPTLILA 234
                 :::.  :  *    ** :.:   * .::.     :: .:       ::*.** :  .

AcsF_Ch          DRYSKPLINLPEDSKAVVAVKKILQNAGII---- 252
CooC1_Ch         SLKGEEIWQSTNPAFVNLHDIYQKLRLEVG---- 254
CooC3_Ch         DLSGVSVFEQAPEIVAIGHSIKERLLEEVKRQNC 268
                 .  .  : : .    .            :
```

Figure A.6: Sequence comparison of AcsF$_{Ch}$ with CooC1$_{Ch}$ and CooC3$_{Ch}$. Sequences were aligned with ClustalW [120]. Dots and stars below the alignment mark similar residues (.), strongly similar residues (:), and identical residues (*). AcsF$_{Ch}$ shares 28.6 % and 28.3 % sequence identity with CooC1$_{Ch}$ and CooC3$_{Ch}$, respectively.

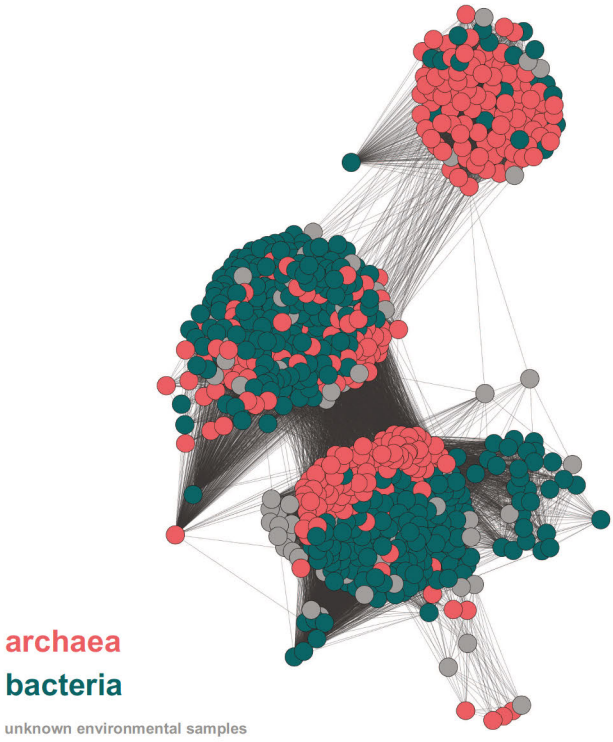


Figure A.7: Distribution of bacterial and archaeal sequences within the sequence similarity network of the family IPRO14433. Nodes in green depict sequences which are derived from bacteria while pink nodes represent sequences derived from archaea. Nodes in grey represent sequences from unknown environmental samples.

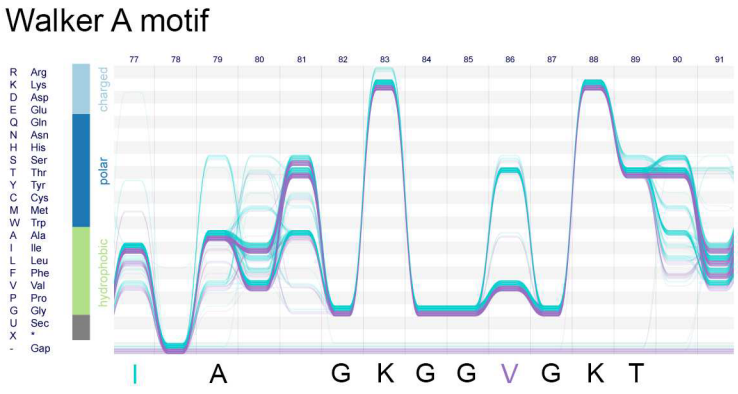


Figure A.8: Sequence bundle alignment of the Walker A motif. Sequences of CooC- and AcsF-type proteins are shown in violet and turquois, respectively. The black single-letter amino acid code is depicted below the alignment, if the residue is conserved in more than 85 % of all sequences. The code is depicted in violet and turquoise, if the residue is conserved in more than 85 % of sequences of the CooC- and AcsF-type subgroup, respectively. Numbers in the top refer to the number of the site in the alignment. Strikingly, the signature lysine residue, thought to be a conserved feature of MinD type proteins (see section 1.4), is replaced by an arginine residue in 12 sequences of the AcsF group (4 %). Sequences which contain an arginine residue instead of the signature lysine are derived from archaeal organisms or from unidentified environmental samples from marine sediment. It is conceivable that arginine functions in a similar way to the signature lysine, as it is also positively charged and has a similar length and flexibility. For 35 sequences the alignment indicates that the Walker A motif is absent. In all likelihood, however, this sequence gap does not actually exist; rather, it is produced by the incorrect assignment of the start codon.

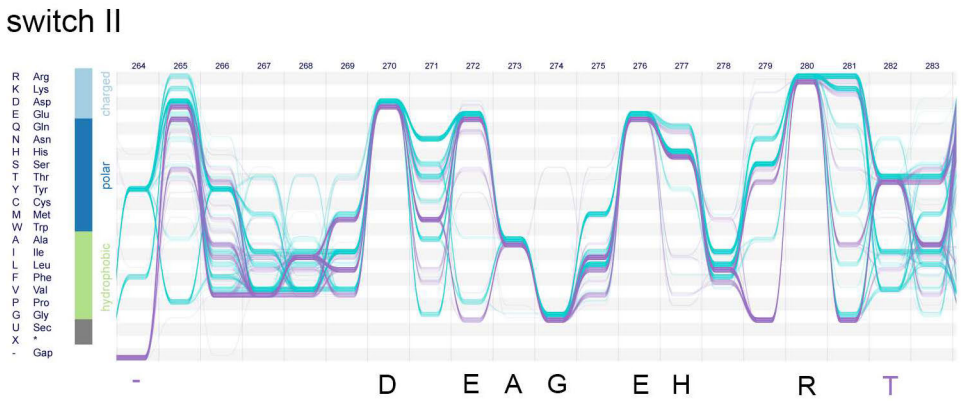


Figure A.9: Sequence bundle alignment of the switch II motif. The color coding is the same as in Fig. A.8.

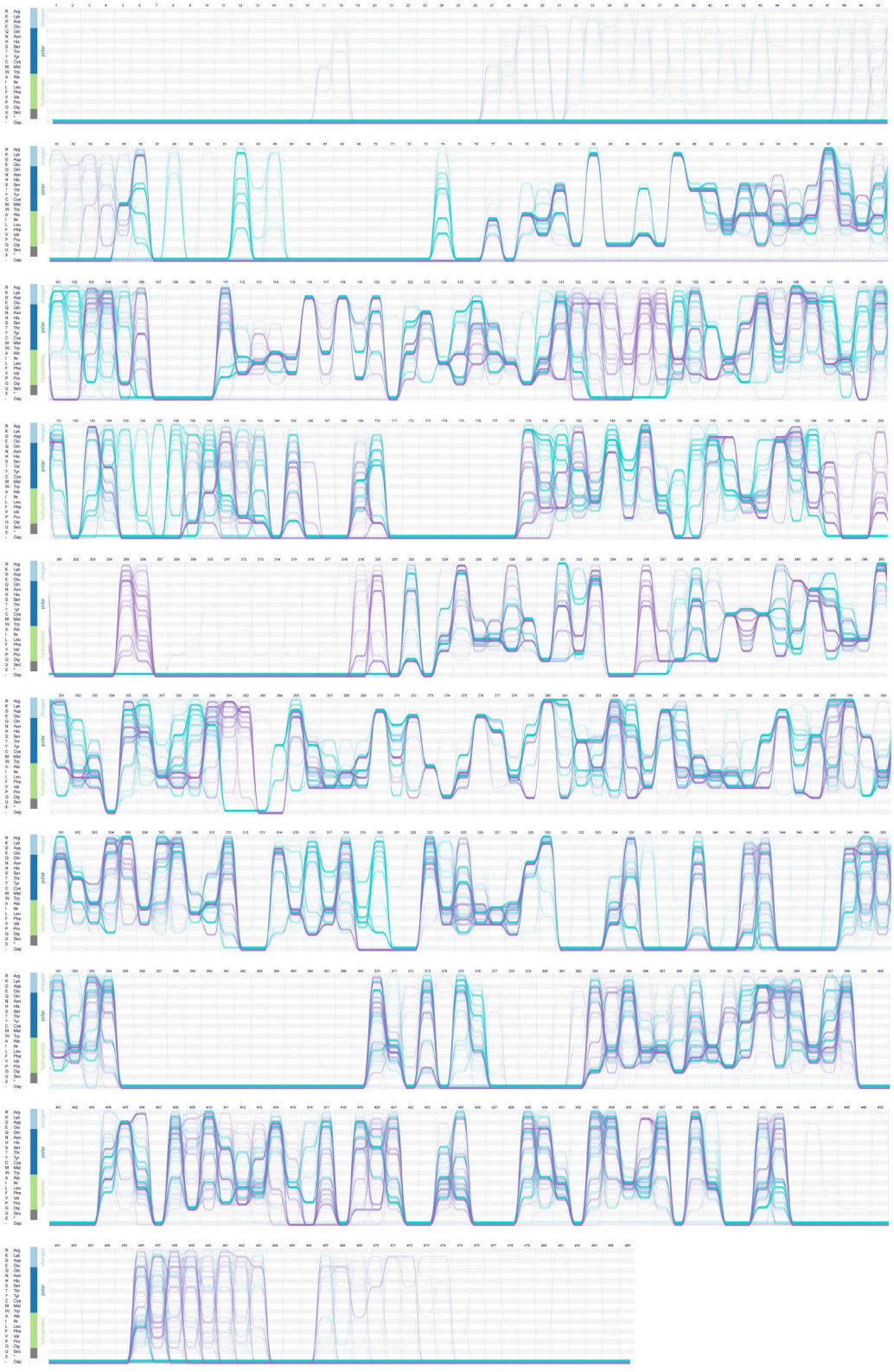


Figure A.10: Complete alignment of AcsF- (turquoise) and CooC-type (violet) proteins.

Table A.2: Data collection and refinement statistics for AcsF$_{Ch}$. Values in parentheses are for the highest resolution shell.

	AcsF$_{Ch}$
Data collection	
Space group	$P\,4_1\,2_1\,2$
Cell dimensions	
a, *b*, *c* (Å)	63.86, 63.86, 228.40
α, β, γ (°)	90, 90, 90
Resolution (Å)	41.99 - 2.65 (2.745 - 2.65)
$I/\sigma(I)$	25.14 (1.93)
$CC_{1/2}$	1.00 (0.72)
Completeness (%)	1.00 (1.00)
Multiplicity	12.9 (12.4)
Refinement	
Resolution (Å)	41.99 - 2.65
No. reflections	14553 (1413)
R_{work}	0.2608 (0.3408)
R_{free}	0.2869 (0.4039)
No. atoms	2002
macromolecules	1946
ligands	37
Protein residues	252
Average B-factor	92.74
macromolecules	92.26
ligands	132.04
solvent	65.92
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.51
Ramachandran statistics	
Ramachandran favored (%)	95
Ramachandran allowed (%)	4.8
Ramachandran outliers (%)	0.4

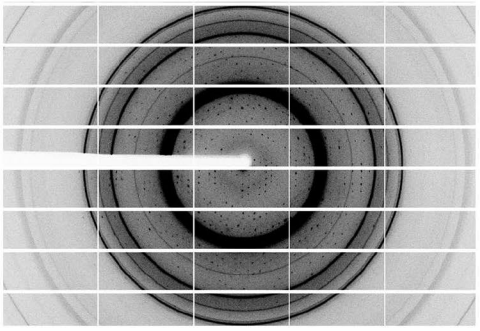
Table A.3: Data collection and refinement statistics for AcsF$_{Af}$. Values in parentheses are for the highest resolution shell.

	monomeric AcsF$_{Af}$	**dimeric AcsF$_{Af}$**
Data collection		
Space group	$C\,2\,2\,2_1$	$C\,2\,2\,2_1$
Cell dimensions		
a, *b*, *c* (Å)	89.95, 105.62, 55.27	110.34, 117.13, 55.79
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	34.24 - 1.7 (1.761 - 1.7)	45.82 - 2.0 (2.071 - 2.0)
$I/\sigma(I)$	12.26 (1.56)	22.04 (2.02)
$CC_{1/2}$	0.998 (0.766)	0.999 (0.811)
Completeness (%)	0.82	1.00 (1.00)
Multiplicity	3.6 (3.7)	6.7 (7.0)
Refinement		
Resolution (Å)	34.24 - 1.7	45.82 - 2.0
No. reflections	24085 (2843)	24815 (2410)
R_{work}	0.1951 (0.2969)	0.1991 (0.3638)
R_{free}	0.2309 (0.3330)	0.2251 (0.4027)
No. atoms	2253	2087
macromolecules	2043	2025
ligands	-	6
Protein residues	258	258
Average B-factor	36.33	65.14
macromolecules	36.36	65.44
ligands	-	75.81
solvent	35.94	53.28
R.m.s. deviations		
Bond lengths (Å)	0.002	0.002
Bond angles (°)	0.50	0.51
Ramachandran statistics		
Ramachandran favored (%)	98	97
Ramachandran allowed (%)	2.3	3.1
Ramachandran outliers (%)	0	0

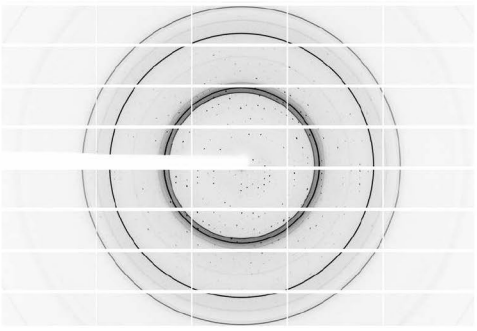
Table A.4: Data collection and refinement statistics for CooC3$_{Ch}$. Values in parentheses are for the highest resolution shell.

	CooC3$_{Ch}$
Data collection	
Space group	$R\,3$
Cell dimensions	
a, *b*, *c* (Å)	194.33, 194.33, 194.33
α, β, γ (°)	90, 90, 120
Resolution (Å)	19.79 - 2.2 (2.279 - 2.2)
$I/\sigma\,(I)$	7.38 (2.27)
$CC_{1/2}$	0.99 (0.92)
Completeness (%)	0.96 (0.96)
Multiplicity	12.9 (12.4)
Refinement	
Resolution (Å)	19.79 - 2.2 (2.279 - 2.2)
No. reflections	30884 (3106)
R_{work}	0.2141 (0.4012)
R_{free}	0.2542 (0.3883)
No. atoms	4020
macromolecules	3830
ligands	28
Protein residues	506
Average B-factor	59.10
macromolecules	59.36
ligands	89.65
solvent	47.57
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.55
Ramachandran statistics	
Ramachandran favored (%)	97
Ramachandran allowed (%)	3
Ramachandran outliers (%)	0.2

A



B



C

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE >= -3.0 AS FUNCTION OF RESOLUTION

RESOLUTION LIMIT	NUMBER OF REFLECTIONS OBSERVED	UNIQUE	POSSIBLE	COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR expected	COMPARED	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano
7.60	1266	370	382	96.9%	1.6%	2.0%	1247	51.52	1.8%	100.0*	-9	0.714	209
5.38	2070	638	641	99.5%	2.4%	2.5%	2017	40.52	2.9%	99.9*	-1	0.775	240
4.39	2869	803	808	99.4%	2.2%	2.3%	2846	46.03	2.5%	99.9*	-23	0.709	410
3.80	3049	837	954	87.7%	2.5%	2.5%	3028	41.07	3.0%	99.9*	-20	0.741	465
3.40	3032	923	1062	86.9%	3.2%	2.8%	2997	33.94	3.9%	99.8*	-11	0.827	371
3.10	3862	1146	1160	98.8%	3.5%	3.3%	3800	30.42	4.1%	99.8*	-5	0.852	474
2.87	4703	1263	1273	99.2%	4.6%	4.5%	4679	24.93	5.4%	99.7*	-12	0.752	735
2.69	5102	1340	1352	99.1%	5.5%	5.7%	5079	20.24	6.4%	99.6*	-5	0.745	813
2.53	5494	1423	1432	99.4%	6.8%	6.9%	5471	17.13	7.9%	99.5*	-5	0.769	881
2.40	5625	1506	1515	99.4%	8.1%	8.5%	5598	13.90	9.5%	99.4*	-1	0.773	838
2.29	5459	1581	1590	99.4%	10.6%	10.9%	5394	10.57	12.6%	98.9*	-2	0.758	604
2.19	2070	587	1673	35.1%	19.0%	17.7%	2046	6.89	22.3%	97.7*	-8	0.729	280
2.11	6451	1716	1722	99.7%	16.6%	17.4%	6419	7.40	19.3%	97.4*	-1	0.756	971
2.03	6423	1728	1792	96.4%	21.0%	19.4%	6395	6.33	24.5%	96.3*	-10	0.726	956
1.96	6906	1815	1845	98.4%	26.9%	28.0%	6878	4.67	31.3%	95.5*	-2	0.724	1051
1.90	1304	356	1936	18.4%	34.2%	32.7%	1292	3.67	39.9%	96.0*	-9	0.682	192
1.84	0	0	1963	0.0%	-99.9%	-99.9%	0	-99.00	-99.9%	0.0	0	0.000	0
1.79	6553	1885	2053	91.8%	46.1%	48.2%	6498	2.43	54.6%	90.2*	0	0.716	715
1.74	7533	2057	2074	99.2%	61.3%	64.1%	7476	1.93	71.7%	82.5*	-1	0.693	1031
1.70	7878	2112	2137	98.8%	71.4%	75.7%	7825	1.68	83.2%	76.1*	-3	0.697	1104
total	87649	24086	29364	82.0%	6.4%	6.5%	86985	14.64	7.6%	99.9*	-5	0.741	12340

Figure A.11: A) and B) Diffraction images of AcsF$_{Af}$ crystals in low and high contrast, respectively. The diffraction pattern of the protein crystal is covered by ice-rings. C) Data integration and scaling statistics of the dataset after the data affected by ice rings had been excluded in the resolution range 3.92 - 3.87 Å, 3.67 - 3.66 Å, 3.45- 3.43 Å, 2.28 - 2.22 Å and 1.95 - 1.84 Å. The affected resolution shells have very low completeness.

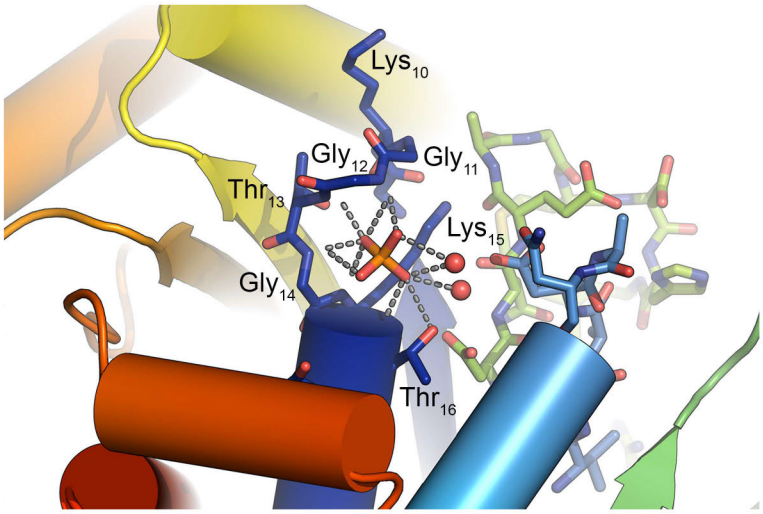


Figure A.12: Phosphate binding site in AcsF$_{Ch}$. Residues of the Walker A motif are shown as blue sticks. Residues of the switch I and II motifs are shown as light blue and light green sticks, respectively. The phosphate molecule forms hydrogen bonds to the protein backbone of Gly12, Thr13 and Gly14 of the Walker A motif and to the hydroxyl group of Thr16.

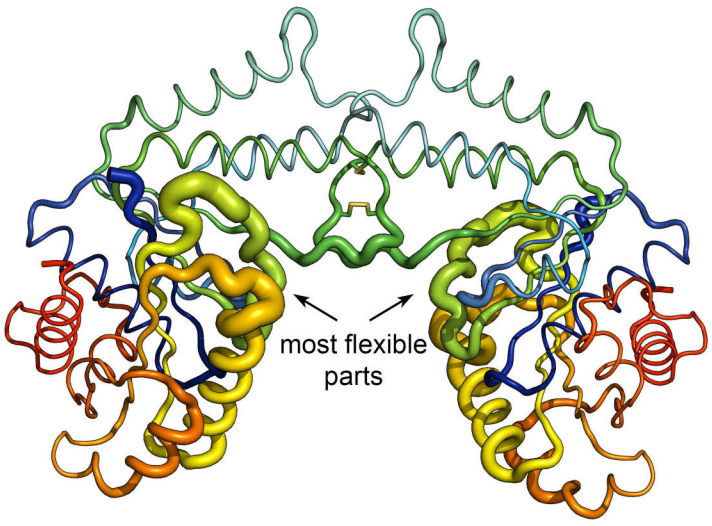


Figure A.13: Sausage representation of AcsF$_{Ch}$.

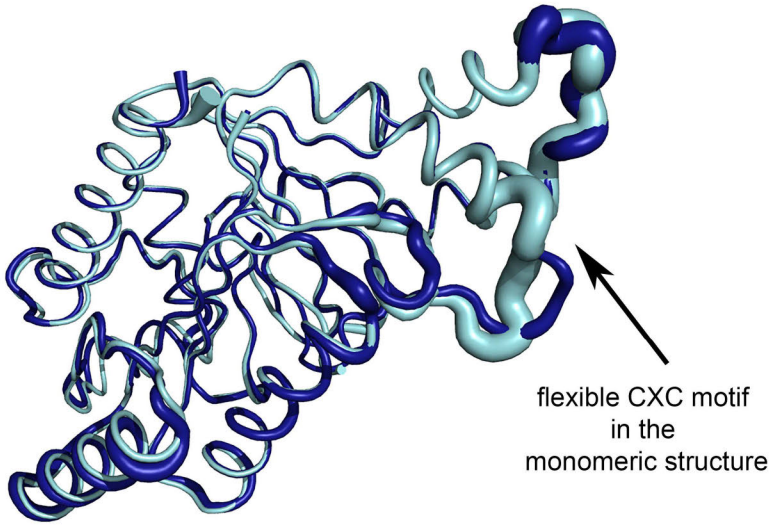


Figure A.14: Sausage representation of monomeric and dimeric AcsF$_{Af}$, which are shown in light and dark blue, respectively.

Table A.5: DALI structural similarity search for AcsF$_{Ch}$.

protein	PDB	Z-score
CooC1	3KJI	22.0
MinD	1HYG	20.5
ParA	3EA0	18.9
Light independent protochlorophyllide reductase	3END	18.5
Nitrogenase iron protein	2AFI	18.5
SOJ	1WCV	17.0
HypB	5AUP	17.0
Mipz	2XJ9	16.6
Get3	3IO3	15.9
Cobrinic acid a,c-diamide synthase	5IF9	15.4
TRC40	3ZQ6	15.2

Table A.6: DALI structural similarity search for AcsF$_{Af}$.

protein	PDB	Z-score
CooC1	3KJI	24.9
MinD	1HYG	23.6
Light independent protochlorophyllide reductase	3END	22.5
Nitrogenase iron protein	2AFI	21.9
ParA	3EA0	21.4
SOJ	1WCV	20.1
Cobrinic acid a,c-diamide synthase	5IF9	18.4
HypB	5AUP	17.9
Mipz	2XJ9	16.6
HypB	5AUP	17.9
Nucleotide-binding protein AF 226A	3KB1	16.6
TRC40	3ZQ6	16.0
Get3	3IBG	15.9
TadZ	3FKQ	15.3
Nucleotide-binding protein AF 2382	2PH1	15.3

Bibliography

[1] A. L. Sommer and C. B. Lipman. Evidence on the indispensable nature of zinc and boron for higher green plants. *Plant Physiology*, 1(3):231, 1926.

[2] A. L. Sommer. Further evidence of the essential nature of zinc for the growth of higher green plants. *Plant Physiology*, 3(2):217, 1928.

[3] J. S. McHargue. Further evidence that small quantities of copper, manganese and zinc are factors in the metabolism of animals. *American Journal of Physiology – Legacy Content*, 77(2):245–255, 1926.

[4] E. Gris. Mémoir relatif à l'action des composés solubles ferrugineux sur la végétation. *Comptes rendus de l'Académie des Sciences*, 17:679, 1843.

[5] R. Bartha and E. J. Ordal. Nickel-dependent chemolithotrophic growth of two *Hydrogenomonas* strains. *Journal of Bacteriology*, 89(4):1015–9, 1965.

[6] D. Bertrand and A. De Wolf. Le nickel, oligoélément dynamique pour les végétaux supérieurs. *Comptes rendus de l'Académie des Sciences*, 265:1053–1055, 1967.

[7] V. Svetlitchnyi, H. Dobbek, W. Meyer-Klaucke, T. Meins, B. Thiele, P. Römer, R. Huber, and O. Meyer. A functional Ni-Ni-[4Fe-4S] cluster in the monomeric acetyl-CoA synthase from *Carboxydothermus hydrogenoformans*. *Proceedings of the National Academy of Science*, 101(2):446–451, 2004.

[8] Y. Dai, P. C. Wensink, and R. H. Abeles. One protein, two enzymes. *Journal of Biological Chemistry*, 274(3):1193–1195, 1999.

[9] G. B. Diekert, E. G. Graf, and R. K. Thauer. Nickel requirement for carbon monoxide dehydrogenase formation in *Clostridium pasteurianum*. *Archives of Microbiology*, 122(1):117–120, 1979.

[10] H. Guldan, R. Sterner, and P. Babinger. Identification and characterization of a bacterial glycerol-1-phosphate dehydrogenase: Ni^{2+}-dependent AraM from *Bacillus subtilis*. *Biochemistry*, 47:7376–7384, 2008.

[11] S. L. Clugston, J. F. J. Barnard, R. Kinach, D. Miedema, R. Ruman, E. Daub, and J. F. Honek. Overproduction and characterization of a dimeric non-zinc glyoxalase I from *Escherichia coli*: Evidence for optimal activation by nickel ions. *Biochemistry*, 37(24):8754–8763, 1998.

[12] A. Ariza, T. J. Vickers, N. Greig, K. A. Armour, M. J. Dixon, I. M. Eggleston, A. H. Fairlamb, and C. S. Bond. Specificity of the trypanothione-dependent *Leishmania major* glyoxalase I: Structure and biochemical comparison with the human enzyme. *Molecular Microbiology*, 59(4):1239–1248, 2006.

[13] B. Desguin, P. Goffin, E. Viaene, M. Kleerebezem, V. Martin-Diaconescu, M. J. Maroney, J.-P. Declercq, P. Soumillion, and P. Hols. Lactate racemase is a nickel-dependent enzyme activated by a widespread maturation system. *Nature Communications*, 5:3615, 2014.

[14] W. L. Ellefson, W. B. Whitman, and R. S. Wolfe. Nickel-containing factor F430: Chromophore of the methylreductase of *Methanobacterium*. *Proceedings of the National Academy of Sciences*, 79(12):3707–3710, 1982.

[15] H. D. Youn, E. J. Kim, J. H. Roe, Y. C. Hah, and S. O. Kang. A novel nickel-containing superoxide dismutase from *Streptomyces* spp. *Biochemical Journal*, 318(3):889–896, 1996.

[16] B. Friedrich, E. Heine, A. Finck, and C. G. Friedrich. Nickel requirement for active hydrogenase formation in *Alcaligenes eutrophus*. *Journal of Bacteriology*, 145(3):1144–1149, 1981.

[17] N. E. Dixon, C. Gazzola, R. L. Blakeley, and B. Zerner. Jack bean urease (EC 3.5.1.5). A metalloenzyme. Simple biological role for nickel? *Journal of the American Chemical Society*, 97(14):4131–4133, 1975.

[18] H. Merkens, R. Kappl, R. P. Jakob, F. X. Schmid, and S. Fetzner. Quercetinase QueD of *Streptomyces* sp. FLA, a monocupin dioxygenase with a preference for nickel and cobalt. *Biochemistry*, 47(46):12185–12196, 2008.

[19] Q. Xu, R. Schwarzenbacher, S. Sri Krishna, D. McMullan, S. Agarwalla, K. Quijano, P. Abdubek, E. Ambing, H. Axelrod, T. Biorac, et al. Crystal structure of acireductone dioxygenase (ARD) from *Mus musculus* at 2.06 Å resolution. *Proteins: Structure, Function, and Bioinformatics*, 64(3):808–813, 2006.

[20] M. M. He, S. L. Clugston, J. F. Honek, and B. W. Matthews. Determination of the structure of *Escherichia coli* glyoxalase I suggests a structural basis for differential metal activation. *Biochemistry*, 39(30):8719–8727, 2000.

[21] J.-H. Jeoung, D. Nianios, S. Fetzner, and H. Dobbek. Quercetin 2, 4-dioxygenase activates dioxygen in a side-on O_2-Ni complex. *Angewandte Chemie International Edition*, 55(10):3281–3284, 2016.

[22] J.-H. Jeoung and H. Dobbek. Carbon dioxide activation at the Ni, Fe-cluster of anaerobic carbon monoxide dehydrogenase. *Science*, 318(5855):1461–1464, 2007.

[23] A. Volbeda, E. Garcin, C. Piras, A. L. de Lacey, V. M. Fernandez, E. C. Hatchikian, M. Frey, and J. C. Fontecilla-Camps. Structure of the [NiFe] hydrogenase active site: Evidence for biologically uncommon Fe ligands. *Journal of the American Chemical Society*, 118(51):12989–12996, 1996.

[24] Y. Zhang, D. Rodionov, M. S. Gelfand, and V. N. Gladyshev. Comparative genomic analyses of nickel, cobalt and vitamin B12 utilization. *BMC Genomics*, 10:78, 2009.

[25] I. A. Berg, D. Kockelkorn, W. H. Ramos-Vera, R. F Say, J. Zarzycki, M. Hügler, B. E. Alber, and G. Fuchs. Autotrophic carbon fixation in archaea. *Nature Reviews Microbiology*, 8(6):447–460, 2010.

[26] V. A. Svetlichny, T. G. Sokolova, M. Gerhardt, M. Ringpfeil, N. A. Kostrikina, and G. A. Zavarzin. *Carboxydothermus hydrogenoformans* gen. nov., sp. nov., a CO-utilizing thermophilic anaerobic bacterium from hydrothermal environments of Kunashir island. *Systematic and Applied Microbiology*, 14(3):254–260, 1991.

[27] K. O. Stetter, G. Lauerer, M. Thomm, and A. Neuner. Isolation of extremely thermophilic sulfate reducers: Evidence for a novel branch of archaebacteria. *Science*, 236:822–824, 1987.

[28] J. Beeder, R. K. Nilsen, J. T. Rosnes, T. Torsvik, and T. Lien. *Archaeoglobus fulgidus* isolated from hot north sea oil field waters. *Applied and Environmental Microbiology*, 60(4):1227–1231, 1994.

[29] J. Vorholt, J. Kunow, K. O. Stetter, and R. K. Thauer. Enzymes and coenzymes of the carbon monoxide dehydrogenase pathway for autotrophic CO_2 fixation in *Archaeoglobus lithotrophicus* and the lack of carbon monoxide dehydrogenase in the heterotrophic *A. profundus*. *Archives of Microbiology*, 163(2):112–118, 1995.

[30] D. Möller-Zinkhan and R. K. Thauer. Anaerobic lactate oxidation to 3 CO_2 by *Archaeoglobus fulgidus* via the carbon monoxide dehydrogenase pathway: Demonstration of the acetyl-CoA carbon-carbon cleavage reaction in cell extracts. *Archives of Microbiology*, 153(3):215–218, 1990.

[31] C. M. Gregg, S. Goetzl, J.-H. Jeoung, and H. Dobbek. AcsF catalyzes the ATP-dependent insertion of Ni into the Ni, Ni-[4Fe4S] cluster of acetyl-CoA synthase. *Journal of Biological Chemistry*, 291(35):18129–18138, 2016.

[32] V. Müller. Energy conservation in acetogenic bacteria. *Applied and Environ mental Microbiology*, 69(11):6345–6353, 2003.

[33] R. K. Thauer, A.-K. Kaster, H. Seedorf, W. Buckel, and R. Hedderich. Methanogenic archaea: Ecologically relevant differences in energy conservation. *Nature Reviews Microbiology*, 6(8):579–591, 2008.

[34] T. I. Doukov, T. M. Iverson, J. Seravalli, S. W. Ragsdale, and C. L. Drennan. A Ni-Fe-Cu center in a bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. *Science*, 298(5593):567–572, 2002.

[35] K. C. Terlesky, M. J. Nelson, and J. G. Ferry. Isolation of an enzyme complex with carbon monoxide dehydrogenase activity containing corrinoid and nickel from acetate-grown *Methanosarcina thermophila*. *Journal of Bacteriology*, 168(3):1053–1058, 1986.

[36] D. A. Grahame. Catalysis of acetyl-CoA cleavage and tetrahydrosarcinapterin methylation by a carbon monoxide dehydrogenase-corrinoid enzyme complex. *Journal of Biological Chemistry*, 266(33):22227–22233, 1991.

[37] W. Gong, B. Hao, Z. Wei, D. J. Ferguson, T. Tallant, J. A. Krzycki, and M. K. Chan. Structure of the $\alpha_2\varepsilon_2$ Ni-dependent CO dehydrogenase component of the *Methanosarcina barkeri* acetyl-CoA decarbonylase/synthase complex. *Proceedings of the National Academy of Sciences*, 105(28):9558–9563, 2008.

[38] W. Shin and P. A. Lindahl. Discovery of a labile nickel ion required for CO/acetyl-CoA exchange activity in the NiFe complex of carbon monoxide dehydrogenase from *Clostridium thermoaceticum*. *Journal of the American Chemical Society*, 114(24):9718–9719, 1992.

[39] C. Darnault, A. Volbeda, E. J. Kim, P. Legrand, X. Vernède, P. A. Lindahl, and J. C. Fontecilla-Camps. Ni-Zn-[Fe4-S4] and Ni-Ni-[Fe4-S4] clusters in closed and open α subunits of acetyl-CoA synthase/carbon monoxide dehydrogenase. *Nature Structural & Molecular Biology*, 10(4):271–279, 2003.

[40] S. Gencic, K. Kelly, S. Ghebreamlak, E. C. Duin, and D. A. Grahame. Different modes of carbon monoxide binding to acetyl-CoA synthase and the role of a conserved phenylalanine in the coordination environment of nickel. *Biochemistry*, 52(10):1705–1716, 2013.

[41] J. Seravalli and S. W. Ragsdale. Pulse-chase studies of the synthesis of acetyl-CoA by carbon monoxide dehydrogenase/acetyl-CoA synthase. *Journal of Biological Chemistry*, 283(13):8384–8394, 2008.

[42] G. Bender and S. W. Ragsdale. Evidence that ferredoxin interfaces with an internal redox shuttle in acetyl-CoA synthase during reductive activation and catalysis. *Biochemistry*, 50(2):276–286, 2010.

[43] S. Gencic, E. C. Duin, and D. A. Grahame. Tight coupling of partial reactions in the acetyl-CoA decarbonylase/synthase (ACDS) multienzyme complex from *Methanosarcina thermophila*. *Journal of Biological Chemistry*, 285(20):15450–15463, 2010.

[44] H. Jones. The cativa process for the manufacture of acetic acid. Iridium catalyst improves productivity in an established industrial process. *Platinum Metals Review*, 44(3):94–105, 2000.

[45] J. F. Roth. The production of acetic acid. *Platinum Metals Review*, 19(1):12–14, 1975.

[46] S. Lutz, A. Jacobi, V. Schlensog, R. Böhm, G. Sawers, and A. Böck. Molecular characterization of an operon (hyp) necessary for the activity of the three hydrogenase isoenzymes in *Escherichia coli*. *Molecular Microbiology*, 5(1):123–135, 1991.

[47] S. Watanabe, D. Sasaki, T. Tominaga, and K. Miki. Structural basis of [NiFe] hydrogenase maturation by Hyp proteins. *Biological Chemistry*, 393(10):1089–1100, 2012.

[48] J. W. Zhang, G. Butland, J. F. Greenblatt, A. Emili, and D. B. Zamble. A role for SlyD in the *Escherichia coli* hydrogenase biosynthetic pathway. *Journal of Biological Chemistry*, 280(6):4360–4366, 2005.

[49] M. R. Leach, J. W. Zhang, and D. B. Zamble. The role of complex formation between the *Escherichia coli* hydrogenase accessory factors HypB and SlyD. *Journal of Biological Chemistry*, 282(22):16177–16186, 2007.

[50] T. Cheng, H. Li, X. Yang, W. Xia, and H. Sun. Interaction of SlyD with HypB of *Helicobacter pylori* facilitates nickel trafficking. *Metallomics*, 5(7):804–807, 2013.

[51] T. Maier, F. Lottspeich, and A. Böck. GTP hydrolysis by HypB is essential for nickel insertion into hydrogenases of *Escherichia coli*. *European Journal of Biochemistry*, 230(1):133–138, 1995.

[52] C. D. Douglas, T. T. Ngu, H. Kaluarachchi, and D. B. Zamble. Metal transfer within the *Escherichia coli* HypB–HypA complex of hydrogenase accessory proteins. *Biochemistry*, 52(35):6030–6039, 2013.

[53] K. C. C. Chung and D. B. Zamble. Protein interactions and localization of the *Escherichia coli* accessory protein HypA during nickel insertion to [NiFe] hydrogenase. *Journal of Biological Chemistry*, 286(50):43081–43090, 2011.

[54] E. Theodoratou, R. Huber, and A. Böck. [NiFe]-hydrogenase maturation endopeptidase: Structure and function. *Biochemical Society Transactions*, 33(1):108–111, 2005.

[55] S. L. Benoit, N. Mehta, M. V. Weinberg, C. Maier, and R. J. Maier. Interaction between the *Helicobacter pylori* accessory proteins HypA and UreE is needed for urease maturation. *Microbiology*, 153(5):1474–1482, 2007.

[56] X. Yang, H. Li, T.-P. Lai, and H. Sun. UreE-UreG complex facilitates nickel transfer and preactivates GTPase of UreG in *Helicobacter pylori*. *Journal of Biological Chemistry*, 290(20):12474–12485, 2015.

[57] A. Soriano and R. P. Hausinger. GTP-dependent activation of urease apoprotein in complex with the UreD, UreF, and UreG accessory proteins. *Proceedings of the National Academy of Sciences*, 96(20):11140–11144, 1999.

[58] Y. H. Fong, H. C. Wong, C. P. Chuck, Y. W. Chen, H. Sun, and K.-B. Wong. Assembly of preactivation complex for urease maturation in *Helicobacter pylori*. *Journal of Biological Chemistry*, 286(50):43241–43249, 2011.

[59] Y. H. Fong, H. C. Wong, M. H. Yuen, P. H. Lau, Y. W. Chen, and K.-B. Wong. Structure of UreG/UreF/UreH complex reveals how urease accessory proteins facilitate maturation of *Helicobacter pylori* urease. *PLoS Biology*, 11(10):e1001678, 2013.

[60] M. A. Farrugia, B. Wang, M. Feig, and R. P. Hausinger. Mutational and computational evidence that a nickel-transfer tunnel in UreD is used for activation of *Klebsiella aerogenes* urease. *Biochemistry*, 54(41):6392–6401, 2015.

[61] R. L. Kerby, P. W. Ludden, and G. P. Roberts. Carbon monoxide-dependent growth of *Rhodospirillum rubrum*. *Journal of Bacteriology*, 177(8):2241–2244, 1995.

[62] R. L. Kerby, P. W. Ludden, and G. P. Roberts. In vivo nickel insertion into the carbon monoxide dehydrogenase of *Rhodospirillum rubrum*: Molecular and physiological characterization of *cooCTJ*. *Journal of Bacteriology*, 179(7):2259–2266, 1997.

[63] R. K. Watt and P. W. Ludden. The identification, purification, and characterization of CooJ. A nickel-binding protein that is CO-regulated with the Ni-containing CO dehydrogenase from *Rhodospirillum rubrum*. *Journal of Biological Chemistry*, 273(16):10019–10025, 1998.

[64] J.-H. Jeoung, T. Giese, M. Grünwald, and H. Dobbek. CooC1 from *Carboxydothermus hydrogenoformans* is a nickel-binding ATPase. *Biochemistry*, 48(48):11505–11513, 2009.

[65] W. B. Jeon, J. Cheng, and P. W. Ludden. Purification and characterization of membrane-associated CooC protein and its functional role in the insertion of nickel into carbon monoxide dehydrogenase from *Rhodospirillum rubrum*. *Journal of Biological Chemistry*, 276(42):38602–38609, 2001.

[66] D. D. Leipe, Y. I. Wolf, E. V. Koonin, and L. Aravind. Classification and evolution of P-loop GTPases and related ATPases. *Journal of Molecular Biology*, 317(1):41–72, 2002.

[67] M. Saraste, P. R. Sibbald, and A. Wittinghofer. The P-loop – a common motif in ATP- and GTP-binding proteins. *Trends in Biochemical Sciences*, 15(11):430–434, 1990.

[68] R. Gasper, A. Scrima, and A. Wittinghofer. Structural insights into HypB, a GTP-binding protein that regulates metal binding. *Journal of Biological Chemistry*, 281(37):27492–27502, 2006.

[69] M. V. Milburn, L. Tong, A. M. deVos, A. Brünger, Z. Yamaizumi, S. Nishimura, and S. H. Kim. Molecular switch for signal transduction: Structural differences between active and inactive forms of protooncogenic ras proteins. *Science*, 247(4945):939, 1990.

[70] C. Fu, J. W. Olson, and R. J. Maier. HypB protein of *Bradyrhizobium japonicum* is a metal-binding GTPase capable of binding 18 divalent nickel ions per dimer. *Proceedings of the National Academy of Sciences*, 92(6):2333–2337, 1995.

[71] J. W. Olson, C. Fu, and R. J. Maier. The HypB protein from *Bradyrhizobium japonicum* can store nickel and is required for the nickel-dependent transcriptional regulation of hydrogenase. *Molecular Microbiology*, 24(1):119–128, 1997.

[72] A. M. Sydor, J. Liu, and D. B. Zamble. Effects of metal on the biochemical properties of *Helicobacter pylori* HypB, a maturation factor of [NiFe]-hydrogenase and urease. *Journal of Bacteriology*, 193(6):1359–1368, 2011.

[73] D. Sasaki, S. Watanabe, R. Matsumi, T. Shoji, A. Yasukochi, K. Tagashira, W. Fukuda, T. Kanai, H. Atomi, T. Imanaka, et al. Identification and structure of a novel archaeal HypB for [NiFe] hydrogenase maturation. *Journal of Molecular Biology*, 425(10):1627–1640, 2013.

[74] S. Watanabe, T. Kawashima, Y. Nishitani, T. Kanai, T. Wada, K. Inaba, H. Atomi, T. Imanaka, and K. Miki. Structural basis of a Ni acquisition cycle for [NiFe] hydrogenase by Ni-metallochaperone HypA and its enhancer. *Proceedings of the National Academy of Sciences*, 112(25):7701–7706, 2015.

[75] T. A. Leonard, P. J. Butler, and J. Löwe. Bacterial chromosome segregation: Structure and DNA binding of the Soj dimer – a conserved biological switch. *The EMBO Journal*, 24(2):270–282, 2005.

[76] J.-H. Jeoung, T. Giese, M. Grünwald, and H. Dobbek. Crystal structure of the ATP-dependent maturation factor of Ni, Fe-containing carbon monoxide dehydrogenases. *Journal of Molecular Biology*, 396(4):1165–1179, 2010.

[77] T. Inoue, K. Takao, Y. Fukuyama, T. Yoshida, and Y. Sako. Over-expression of carbon monoxide dehydrogenase-I with an accessory protein co-expression: A key enzyme for carbon dioxide reduction. *Bioscience, Biotechnology, and Biochemistry*, 78(4):582–587, 2014.

[78] H.-K. Loke and P. A. Lindahl. Identification and preliminary characterization of AcsF, a putative Ni-insertase used in the biosynthesis of acetyl-CoA synthase from *Clostridium thermoaceticum*. *Journal of Inorganic Biochemistry*, 93(1):33–40, 2003.

[79] R. Apweiler, T. K. Attwood, A. Bairoch, A. Bateman, E. Birney, M. Biswas, P. Bucher, L. Cerutti, F. Corpet, M. D. R. Croning, et al. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Research*, 29(1):37–40, 2001.

[80] A. J. Zehnder and K. Wuhrmann. Titanium (III) citrate as a nontoxic oxidation-reduction buffering system for the culture of obligate anaerobes. *Science*, 194(4270):1165–1166, 1976.

[81] H. E. Klock, E. J. Koesema, M. W. Knuth, and S. A. Lesley. Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts. *Proteins: Structure, Function, and Bioinformatics*, 71(2):982–994, 2008.

[82] H. E. Klock and S. A. Lesley. The polymerase incomplete primer extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis. *High Throughput Protein Expression and Purification: Methods and Protocols*, 498:91–103, 2009.

[83] C. Papworth, J. C. Bauer, J. Braman, and D. A. Wright. Site-directed mutagenesis in one day with > 80% efficiency. *Strategies*, 9(4):3–4, 1996.

[84] W. Wang and B. Malcolm. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange site-directed mutagenesis. *Biotechniques*, 26(4):680–682, 1999.

[85] H. Inoue, H. Nojima, and H. Okayama. High efficiency transformation of *Escherichia coli* with plasmids. *Gene*, 96(1):23–28, 1990.

[86] D. W. Cleveland, S. G. Fischer, M. W. Kirschner, and U. K. Laemmli. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. *Journal of Biological Chemistry*, 252(3):1102–1106, 1977.

[87] X. Zou, D. R. Evans, and K. L. Brown. Efficient and convenient method for axial nucleotide removal from vitamin B12 and its derivatives. *Inorganic Chemistry*, 34(7):1634–1635, 1995.

[88] W. D. Robertson, A. M. Bovell, and K. Warncke. Cobinamide production of hydrogen in a homogeneous aqueous photochemical system, and assembly and photoreduction in a $(\beta\alpha)_8$ protein. *Journal of Biological Inorganic Chemistry*, 18(6):701–713, 2013.

[89] J. E. Lindsley. Use of a real-time, coupled assay to measure the ATPase activity of DNA topoisomerase II. *DNA Topoisomerase Protocols*, 95:57–64, 2001.

[90] P. A. Lanzetta, L. J. Alvarez, P. S. Reinach, and O. A. Candia. An improved assay for nanomole amounts of inorganic phosphate. *Analytical Biochemistry*, 100(1):95–97, 1979.

[91] online available at http://www.gelanalyzer.com/.

[92] S. Keller, C. Vargas, H. Zhao, G. Piszczek, C. A. Brautigam, and P. Schuck. High-precision isothermal titration calorimetry with automated peak-shape analysis. *Analytical Chemistry*, 84(11):5066–5073, 2012.

[93] T. Wiseman, S. Williston, J. F. Brandts, and L.-N. Lin. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry*, 179(1):131–137, 1989.

[94] K. Katoh and D. M. Standley. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4):772–780, 2013.

[95] K. Katoh, K. Misawa, K. Kuma, and T. Miyata. MAFFT: A novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Research*, 30(14):3059–3066, 2002.

[96] J. Felsenstein. PHYLIP: Phylogeny inference package (Version 3.2). *Cladistics*, 5:164–166, 1989.

[97] M. E. Smoot, K. Ono, J. Ruscheinski, P.-L. Wang, and T. Ideker. Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics*, 27(3):431–432, 2011.

[98] R. F. Schwarz, A. U. Tamuri, M. Kultys, J. King, J. Godwin, A. M. Florescu, J. Schultz, and N. Goldman. ALVIS: Interactive non-aggregative visualization and explorative analysis of multiple sequence alignments. *Nucleic Acids Research*, 44(8):e77, 2016.

[99] M. Gerlach, U. Mueller, and M. S. Weiss. The MX beamlines BL14.1-3 at BESSY II. *Journal of Large-Scale Research Facilities JLSRF*, 2:47, 2016.

[100] A. G. W. Leslie and H. R. Powell. Processing diffraction data with MOSFLM. *Evolving methods for macromolecular crystallography*, 245:41–51, 2007.

[101] W. Kabsch. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D: Biological Crystallography*, 66(2):133–144, 2010.

[102] M. Krug, M. S. Weiss, U. Heinemann, and U. Mueller. XDSAPP: A graphical user interface for the convenient processing of diffraction data using XDS. *Journal of Applied Crystallography*, 45(3):568–572, 2012.

[103] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read. *Phaser* crystallographic software. *Journal of Applied Crystallography*, 40(4):658–674, 2007.

[104] G. Bunkóczi and R. J. Read. Improvement of molecular-replacement models with *Sculptor*. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):303–312, 2011.

[105] T. C. Terwilliger, R. W. Grosse-Kunstleve, P. V. Afonine, N. W. Moriarty, P. H. Zwart, L.-W. Hung, R. J. Read, and P. D. Adams. Iterative model building, structure refinement and density modification with the *PHENIX AutoBuild* wizard. *Acta Crystallographica Section D: Biological Crystallography*, 64(1):61–69, 2008.

[106] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan. Features and development of *Coot*. *Acta Crystallographica Section D: Biological Crystallography*, 66(4):486–501, 2010.

[107] P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, and P. D. Adams. Towards automated crystallographic structure refinement with *phenix.refine*. *Acta Crystallographica Section D: Biological Crystallography*, 68(4):352–367, 2012.

[108] Schrödinger, LLC. The PyMOL molecular graphics system. 2015.

[109] A. Stivala, M. Wybrow, A. Wirth, J. C. Whisstock, and P. J. Stuckey. Automatic generation of protein structure cartoons with Pro-origami. *Bioinformatics*, 27(23):3315–3316, 2011.

[110] W. Xia, H. Li, X. Yang, K.-B. Wong, and H. Sun. Metallo-GTPase HypB from *Helicobacter pylori* and its interaction with nickel chaperone protein HypA. *Journal of Biological Chemistry*, 287(9):6753–6763, 2012.

[111] I. Hayashi, T. Oyama, and K. Morikawa. Structural and functional studies of MinD ATPase: Implications for the molecular recognition of the bacterial cell division apparatus. *The EMBO Journal*, 20(8):1819–1828, 2001.

[112] N. Sakai, M. Yao, H. Itou, N. Watanabe, F. Yumoto, M. Tanokura, and I. Tanaka. The three-dimensional structure of septum site-determining protein MinD from *Pyrococcus horikoshii* OT3 in complex with Mg-ADP. *Structure*, 9(9):817–826, 2001.

[113] U. Tokumoto and Y. Takahashi. Genetic analysis of the *isc* operon in *Escherichia coli* involved in the biogenesis of cellular iron-sulfur proteins. *Journal of Biochemistry*, 130(1):63–71, 2001.

[114] H.-K. Loke, X. Tan, and P. A. Lindahl. Genetic construction of truncated and chimeric metalloproteins derived from the α subunit of acetyl-CoA synthase from *Clostridium thermoaceticum*. *Journal of the American Chemical Society*, 124(29):8667–8672, 2002.

[115] S. J. George, J. Seravalli, and S. W. Ragsdale. EPR and infrared spectroscopic evidence that a kinetically competent paramagnetic intermediate is formed when acetyl-coenzyme A synthase reacts with CO. *Journal of the American Chemical Society*, 127(39):13500–13501, 2005.

[116] L. Holm and P. Rosenström. Dali server: Conservation mapping in 3D. *Nucleic Acids Research*, 38(suppl 2):W545–W549, 2010.

[117] J. A. Denny and M. Y. Darensbourg. Metallodithiolates as ligands in coordination, bioinorganic, and organometallic chemistry. *Chemical Reviews*, 115(11):5248–5273, 2015.

[118] P. V. Rao, S. Bhaduri, J. Jiang, and R. H. Holm. Sulfur bridging interactions of cis-planar Ni^{II}-S_2N_2 coordination units with nickel(II), copper(I, II), zinc(II), and mercury(II): A library of bridging modes, including $Ni^{II}(\mu_2\text{-}SR)_2M^{I,\,II}$ rhombs. *Inorganic Chemistry*, 43(19):5833–5849, 2004.

[119] Y. Ye and A. Godzik. Flexible structure alignment by chaining aligned fragment pairs allowing twists. *Bioinformatics*, 19(suppl 2):ii246–ii255, 2003.

[120] M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21):2947–2948, 2007.

List of Figures

1.1 Overview of the active sites of nickel-containing enzymes 3
1.2 The reductive acetyl-CoA pathway in *C. hydrogenoformans* 6
1.3 Crystal structures of ACS 9
1.4 Two proposed mechanisms of ACS 10
1.5 Proposed pathway of the maturation of [NiFe]-hydrogenase 12
1.6 Proposed pathway of nickel delivery into *H. pylori* urease 13
1.7 Crystal structures of HypB and UreG 16

3.1 Purification of $AcsF_{Ch}$ 36
3.2 Purification of $AcsF_{Af}$ 37
3.3 Purification of ACS 38
3.4 UV-Vis spectrum of $AcsF_{Ch}$ and ACS 39
3.5 Titration of $AcsF_{Ch}$ with $NiCl_2$ 40
3.6 Nucleotide dependent dimerization of $AcsF_{Ch}$ and K10A-$AcsF_{Ch}$ 42
3.7 Complex formation of apoACS with $AcsF_{Ch}$ 45
3.8 Complex formation behaviour of apoACS with CooC-type proteins and $AcsF_{Ch}$ variants 46
3.9 ACS activation by $AcsF_{Ch}$, Mg-ATP and $NiCl_2$ 47
3.10 ACS activation by variants of $AcsF_{Ch}$ 48
3.11 ITC of ACS and ACS-$AcsF_{Ch}$ with $NiCl_2$ 50
3.12 Stepwise activation of ACS 51
3.13 Subgroups of the CooC family of proteins 52
3.14 Sequence bundle alignment of the switch I motif 53

3.15 Sequence bundle alignment of the CXC motif 53
3.16 Sequential differences at site 140 54
3.17 Protein crystals of $AcsF_{Ch}$ 55
3.18 Protein crystals of $AcsF_{Af}$ 56
3.19 Green-coloured $CooC3_{Ch}$ crystals 57
3.20 Crystal structure of $AcsF_{Ch}$ 59
3.21 Positions of the structural motifs relative to each other 60
3.22 Crystal structure and topology diagram of $AcsF_{Af}$ 62
3.23 Zinc binding site of $AcsF_{Af}$ 62
3.24 Nickel binding site of $CooC3_{Ch}$ 64
3.25 Structural comparison of AcsF- and CooC-type proteins 66
3.26 Comparison of helix α3 of $AcsF_{Af}$ and the equivalent positions in $CooC1_{Ch}$ 67
3.27 Superimposed CXC motif of AcsF and CooC proteins 68

4.1 Preliminary model of ACS activation by $AcsF_{Ch}$ 72

A.1 $AcsF_{Af}$ gene with restriction sites 75
A.2 Sequence and structure of the search model “short-CooC1” 76
A.3 Sequence and structure of the search model “short-AcsF” 77
A.4 Expression level of the gene encoding $AcsF_{Ch}$ 78
A.5 Calibration curve of the superdex S200 column 79
A.6 Sequence comparison of $AcsF_{Ch}$ with $CooC1_{Ch}$ and $CooC3_{Ch}$ 80
A.7 Sequence similarity network of the family IPRO14433 showing bacterial and archaeal proteins 81
A.8 Sequence bundle alignment of the Walker A motif 82
A.9 Sequence bundle alignment of the switch II motif 83
A.10 Complete alignment of AcsF- and CooC-type proteins 84
A.11 Diffraction images of $AcsF_{Af}$ crystals with ice rings 88
A.12 Phosphate binding site in $AcsF_{Ch}$ 89

A.13 Sausage representation of AcsF$_{Ch}$ 89
A.14 Sausage representation of monomeric and dimeric AcsF$_{Af}$ 90

List of Tables

1.1 Nickel containing enzymes 4
1.2 Autotrophic carbon fixation pathways 7

2.1 Composition of cultivation media 20
2.2 PCR protocol for PIPE cloning 22
2.3 PCR protocol for quikchange mutagenesis 22

3.1 ATPase activity of AcsF$_{Ch}$ 43
3.2 Metal analysis of ACS samples by ICP-OES 50
3.3 FATCAT structural alignment 69

A.1 PCR primers used for cloning 75
A.2 Data collection and refinement statistics for AcsF$_{Ch}$ 85
A.3 Data collection and refinement statistics for AcsF$_{Af}$ 86
A.4 Data collection and refinement statistics for CooC3$_{Ch}$ 87
A.5 DALI structural similarity search for AcsF$_{Ch}$ 91
A.6 DALI structural similarity search for AcsF$_{Af}$ 91

List of Abbreviations

Å	Ångstrom
ACDS	acetyl-CoA decarbonylase/synthase
ACS	acetyl-CoA synthase
ADP	adenosine diphosphate
A. fulgidus	*Archaeoglobus fulgidus*
ATP	adenosine triphosphate
ATPase	ATP hydrolase
APS	ammonium persulfate
BuOH	butanol
Cbi	cobinamide
C. hydrogenoformans	*Carboxydothermus hyrogenoformans*
CoA	coenzyme A
CODH	carbon monoxide dehydrogenase
CoFeSP	corrinoid iron sulfur protein
CoM	coenzyme M
CV	column volume
D. vulgaris	*Desulfovibrio vulgaris*
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
$F_{420}H_2$	cofactor F420, reduced
F_{430}	cofactor F430
GDP	guanosine diphosphate
GTP	guanosine triphosphate

GTPase	GTP hydrolase
H. pylori	*Helicobacter pylori*
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
LB	lysogeny broth
LDH	lactate dehydrogenase
MeCbi	methylcobinamide
MeTr	methyl transferase
mTB	modified terrific broth
M. thermoacetica	*Moorella thermoacetica*
NADH	nicotinamide adenine dinucleotide, reduced
Ni_d	distal nickel
Ni_p	proximal nickel
NDP	nucleoside diphosphate
NTP	nucleoside triphosphate
NTPase	NTP hydrolase
OD_{600}	optical density at 600 nm
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
P_i	inorganic phosphate
PIPE	polymerase incomplete primer extension
PK	pyruvate kinase
RACo	reductive activator of CoFeSP
R. rubrum	*Rhodospirillium rubrum*
SDS	sodium dodecyl sulfate
SIMIBI	signal recognition particle, MinD and BioD

SOC	super optimal broth with catabolic repression
SOB	super optimal broth
TEMED	N,N,N',N'-tetramethylethylenediamine
TRAFAC	translation factor related
TAE	Tris-acetate-EDTA
TCEP	tris(2-carboxyethyl)phosphine
TEV protease	tabacco etch virus protease
Tris	tris(hydroxymethyl)aminomethane
UV-Vis	ultraviolet-visible
v/v	volume / volume
w/v	weight / volume

Declaration

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine weiteren als die angegebenen Hilfsmittel verwendet habe.

I hereby declare that I have prepared this work independently, using only references and resources, which are marked as such.

Berlin, den 27.07.2016 Christina Gregg

Acknowledgement

Firstly, I would like to thank Prof. Holger Dobbek for his guidance and support over the past four years, and for giving me the opportunity to work in his research group.

I am deeply indebted to Dr. Jae-Hun Jeoung, who taught me so much about CooC proteins, without which my research would not have been possible. He always made himself available when his assistance was needed.

I would like to thank Dr. Sebastian Götzl for introducing me to ACS, and Yulia Ilina for helpful discussions on this topic.

Many thanks to Dr. Martin Bommer, who kindly provided vectors for cloning and thanks to Dr. Tobias Werther for providing valuable input to the ITC experiments.

I am grateful to Jasmin Kurtzke from Silke Leimkühler's lab for analyzing the metal content of my samples.

Aside from these individual contributions, every one of my colleagues over the past four years has been an absolute pleasure to work with. Many thanks to my labmates: Berta, Brinda, Friederike, Gisa, Jae-Hun, Jochen, Lilith, Martin, Rainer, Sabine, Sandra, Sebastian, Tobias und Yulia. Many thanks also to the students who worked in our lab during this time: Anastasia, Cornelia, Daniel, Florian, Lydia, Noam und Peer.

I am especially grateful to Rainer Dietrich for his tireless efforts to keep our lab running smoothly.

Many thanks to Barbara Franke for taking care of administrative tasks.

Finally, I would like to thank my family for their constant encouragement and optimism during my studies. A special thanks to my husband Tim, who has supported me in so many ways.